Massey University Library
Thesis Copyright Form

Title of thesis:

(1) (a) I give permission for my thesis to be made available to readers in Massey University Library under conditions determined by the Librarian.

(b) I do not wish my thesis to be made available to readers without my written consent for ... months.

(2) (a) I agree that my thesis, or a copy, may be sent to another institution under conditions determined by the Librarian.

(b) I do not wish my thesis, or a copy, to be sent to another institution without my written consent for ... months.

(3) (a) I agree that my thesis may be copied for Library use.

(b) I do not wish my thesis to be copied for Library use for ... months.

Signed Ana Estela Antunes da Silva

Date 25.02.94

NAME AND ADDRESS DATE

The AIM Methodology

Analysis, Interpretation And Modelling: A Methodology For The Role Of The Knowledge Analyst In Knowledge Acquisition.

A thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Computer Science at Massey University.

Ana Estela Antunes da Silva

February, 1994.

Abstract

A methodology for the analysis, interpretation and modelling of knowledge is developed in this thesis - the AIM methodology. This methodology can be used by a knowledge engineer who has had no involvement with the elicitation process. The knowledge engineer who carries out the processes of the AIM methodology is called the knowledge analyst.

As a general description, the AIM methodology presents techniques together with their advantages and pitfalls in the processes of analysis, interpretation and modelling. The process of analysis consists of the refinement of the elicited knowledge. During this process, techniques such as the discourse analysis are used. The interpretation and modelling processes aim at representing the domain elements obtained during the analysis process. Models such as the model of expertise of the KADS methodology are used during the interpretation and modelling processes.

This research also aims at defining the activities involved in the knowledge acquisition phase in the development of knowledge-based systems. In this thesis, knowledge acquisition is considered as a set of processes which comprises not only the elicitation of knowledge, but also the refinement of the knowledge elicited through its analysis, interpretation and modelling.

Acknowledgements

Thanks are due to my supervisors, Mrs. Elizabeth Kemp and Mrs. Elisabeth Todd for their invaluable work and encouragement throughout this thesis.

Thanks to Mr. David Gray for the help with the data set used in this thesis.

Ministry of Foreign Affairs and Trade is thanked for the financial support of this study.

For all assistance, specially during my first year at Massey University, I thank Professor Mark Apperley.

Finally, thanks to my husband, Sila for his encouragement and patience.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Knowledge-based systems development has been the subject of many research studies for several years (Guida and Tasso, 1989; Bramer, 1987;Jackson, 1990). Knowledge-based systems can be distinguished from more conventional systems because they use representations of human knowledge instead of algorithmic approaches. They solve problems by heuristics or rules of thumb simulating human reasoning. They should give solutions to complex problems and/or give advice on difficult situations that an expert would do if facing the same circumstances.

One of the most important phases in the development of knowledge-based systems is the knowledge acquisition process. This is often described as a difficult and time consuming task (Motta *et al.,* 1989a; Aamodt, 1990; Althoff *et al,* 1990). Factors such as the difference between the vocabulary used by experts and knowledge engineers, and the gap between the knowledge obtained from experts and the representations in the knowledge base are identified as some of these difficulties. Several methodologies for building knowledge-based systems have been considered in the literature (Motta *et al.*, 1989a; Karbach *et al.,* 1989; Motta *et al.,* 1989b; Wielinga *et al.,* 1991). Although none of them has been identified as the best development solution, the KADS methodology promises to succeed in several areas where other approaches have failed. Knowledge acquisition is one of these areas.

The methodology used in this research aims to position knowledge acquisition as a phase within the development of knowledge-based systems as well as to establish its main activities. This research considers knowledge acquisition as the phase which contains the processes of elicitation, analysis, interpretation and modelling of knowledge. Three of these processes are described in this methodology: **a**nalysis, **i**nterpretation and **m**odelling. Together they comprise the AIM methodology.

The AIM methodology is mainly based on analysis techniques and the model of expertise of the KADS methodology. The research data set is also an important part of this study. The domain data set used to evaluate the AIM methodology was obtained in the first phase of a research project at Massey University. The domain area concerns New Zealand dairy production during the summer-autumn period and involves the main decisions dairy farmers have to make to maximise their profits and the problems they face when making such decisions.

The AIM methodology also presents two different functions for the knowledge engineer. According to Hart (1992), to perform the role of a knowledge engineer a person should possess such characteristics as good communication skills, tact and diplomacy, logicality, technical knowledge and domain knowledge. Having a knowledge engineer with a background in the domain is a crucial point. It is very important that the expert does not think the knowledge engineer is too inexperienced or unclear about the domain during the elicitation process. However, it is difficult for the knowledge engineer to gain a good understanding of a completely new area in a short period of time and talk to the expert at the same level of knowledge. This has given rise to the question of who should be the knowledge engineer, a person with a technical background or a person with domain expertise.

To address this issue, this methodology considers two different roles for the knowledge engineer. The first role is responsible for the knowledge elicitation process and is directly involved with the expert. This person is called the knowledge elicitor. In this research, two participants of the dairy project played the role of the knowledge elicitor. The second role of knowledge engineer is called the knowledge analyst. This person is responsible for all the processes of the AIM methodology: analysis, interpretation and modelling. The knowledge analyst is not directly involved with the expert. The author of this thesis played the role of the knowledge analyst. Although considering the work of the knowledge elicitors as the input data set, the processes of this methodology are designed to be performed by the knowledge analyst. A similar view about different roles for the knowledge engineer is presented by Brule' & Blount (1989). The authors mention the advantage of working as a team which contains two knowledge engineers who would assist each other in the elicitation and analysis processes.

In the second chapter the supporting literature is described emphasising the importance of the knowledge acquisition phase and the different definitions given to this process. This chapter also focuses on the interchangeable way in which knowledge acquisition and knowledge elicitation are mentioned in the literature and explains the modelling approach of the KADS methodology. Chapter 3 describes the structure of the dairy project and its suitability for a knowledge-based system. In chapter 4, the AIM methodology developed in this research is explained. The chapter describes the structure of the AIM methodology and how the techniques and the model of expertise were used. In chapter five, the applicability of the AIM methodology is analyzed through a case study. A discussion of the results of the application of the methodology is also provided. Chapter six contains the conclusion drawn from the research and the future work that can be done.

Chapter 2: Literature Review

2.1 Introduction

The concepts of knowledge acquisition and knowledge elicitation are often used synonymously in the literature without being formally defined. It is necessary here to differentiate between these two concepts and determine a proper definition for each of them. This involves analyzing the views of several authors about knowledge acquisition and knowledge elicitation and how these two procedures are applied in the context of knowledge-based systems.

The need for a methodology for building knowledge-based systems is also emphasised in the review. Such a methodology should contain steps such as knowledge acquisition and knowledge elicitation as well as other procedures and tools which support the development of a knowledge-based system. The emphasis here will be on the analysis phase rather than the design, but the result of such analysis should be the direct input to the design phase, thus bridging, the gap between the expertise elicited from the expert and the knowledge base itself.

In the context of methodologies, the KADS methodology appears to be one of the most promising. Two main concepts in the KADS methodology are responsible for this: the model of expertise and the library of problem-solving tasks. Other methodologies for building knowledge-based systems will also be described in order to compare their performances with the KADS methodology.

For the purpose of this review the KADS methodology will be explained in terms of its models. Some knowledge elicitation techniques will be taken into account apart from the KADS methodology since it does not provide a set of knowledge elicitation techniques of its own.

2.2 Knowledge acquisition and elicitation

One of the most important phases in the development of knowledge-based systems is the knowledge acquisition phase. As mentioned by Wielinga et al. (1987) knowledge acquisition is seen as difficult and considered to be the major bottleneck in knowledge-based systems development. A formal definition for knowledge acquisition, however, does not exist, and the concept is frequently associated with another important phase in knowledge-based systems development - knowledge elicitation.

In fact, some authors do not differentiate between knowledge acquisition and knowledge elicitation. In Welbank (1990) the title claimed to provide an analysis of knowledge acquisition methods and the techniques mentioned in the article were addressed as knowledge elicitation techniques. In order to establish a formal meaning for these concepts some definitions of both terms and how they are used in knowledge-based systems development will be examined.

According to Bylander & Chandrasekaran (1987), knowledge acquisition is "the process that extracts knowledge from a source and incorporates it into a knowledge-based system that solves some problem". This is rather a broad definition, which seems to cover most of the process which has been called knowledge acquisition process. A source may be an expert, a domain, a textbook etc and because the way the knowledge is extracted is not mentioned, this definition gives the option of choosing between either a knowledge elicitor or a piece of software to acquire the knowledge from the expert. The problem with this definition seems to be the assumption that a knowledge-based system incorporates the expert's knowledge, in the sense that the expert system should replicate the job of the expert. However, this problem goes beyond the limits of this review.

The fact that knowledge may be acquired by a knowledge engineer or by software leads to another characteristic of the knowledge acquisition process, that is that knowledge may be extracted from the expert in different means. According to Hayes-Roth *et al.* (1983), knowledge acquisition can be carried out in a number of different ways. In Wielinga & Breuker (1987), these ways were called Modalities of Knowledge Acquisition.

For the purpose of this review, knowledge acquisition will be considered as the process carried out by the knowledge engineer, since no software was involved in the acquisition of the farmer's knowledge. Paradigms of knowledge acquisition are discussed in Breuker *et al.* (1987). The authors mention the fact that knowledge acquisition has to be seen as a modelling metaphor, rather than a metaphor of extraction or mining.

Another definition of knowledge acquisition is given by Neale (1990). In his view knowledge acquisition is a modelling process which consists of two stages. The first stage involves building an outline theory of the domain, and the second stage aims to identify and model the knowledge the expert uses. The first stage is called the competence model and the latter the performance model. This way of seeing knowledge acquisition as a modelling (sometimes called constructive) process has been mentioned in the literature as a new way of visualizing knowledge acquisition. As described by Wielinga et al. (1991) knowledge acquisition is the process of building a computational model of desired behaviour observed or specified in terms of real-world phenomena. The use of models should be a means of coping with the complexity of the process.

It seems that this point-of-view is consistent with the fact that knowledge acquisition can be seen as models of development which have cycles across stages rather than a linear view of the life cycle of a system. This opinion is endorsed by Cordingley (1989) who mentions that the previously dominant linear view of a system's life cycle is being superseded by the modelling process. So, according to these later views, knowledge acquisition is better defined as a cyclical process divided into sections which are sometimes called models.

As a cyclical process, knowledge acquisition has been divided into parts, from collection of data, passing through interpretation and analysis, to modelling. In such definitions it is common to find knowledge elicitation as part of the cyclical process of knowledge acquisition. An example of knowledge elicitation as a part of the knowledge acquisition process is the definition that knowledge acquisition involves **eliciting** the knowledge in an informal form, interpreting the elicited data using some conceptual framework, and formalising the conceptualisations in such a way that the program can use the knowledge (Wielinga *et al.*, 1991). This fact is also mentioned in Kidd (1987) where knowledge acquisition involves **eliciting**, analyzing, and interpreting the knowledge that a human expert uses when solving a particular problem and then transforming this knowledge into a suitable machine representation.

Several authors (Motta, 1989b; Wielinga *et al.,* 1986) share this view that knowledge elicitation is considered an activity in the knowledge acquisition process. A different opinion is given by Cordingley (1989) who proposed a working definition of knowledge elicitation which is detached from the knowledge acquisition process, as those activities undertaken by a person the knowledge elicitor, to obtain material from relevant sources, analyze and interpret the material, and put it in a pre-coded form. A similar opinion of knowledge elicitation is given by Jackson (1990) who says that knowledge elicitation consists of the transfer of potential problem-solving expertise from some knowledge source to a program, accomplished by a series of interviews between a domain expert and a knowledge engineer who then writes a computer program representing the knowledge.

Although the two opinions mentioned above seem to consider knowledge elicitation as an independent procedure, it is less common to find in the literature a formal definition of knowledge elicitation which is not associated with knowledge acquisition. For the purpose of this review, *knowledge elicitation will be seen as an activity in the knowledge acquisition phase.* This activity will consist of obtaining the knowledge from the expert through elicitation techniques such as interviews, and reporting it in a form which both: the knowledge engineer and the expert could understand. This report does not require a formal representation of the knowledge obtained, such as frames or rules (which would be necessary later for the construction of the system and might be the result of an interview). And knowledge acquisition will be considered as *the phase containing the processes of collecting (eliciting), analyzing, interpreting and modelling the data.* Also, knowledge acquisition will be taken not as a linear process but, as an iterative process which should be flexible enough to redo its cycles if necessary.

2.3 Elicitation techniques

As was mentioned in session 2.1, knowledge elicitation is not always well defined and not all authors agree on its meaning, but most of literature agrees on a common set of knowledge elicitation techniques. Knowledge elicitation techniques play a very important role in the construction of the knowledge-based system, being the connection between the knowledge engineer and the expert.

Although these techniques are used with the same purpose, which is to elicit knowledge from the expert, they vary depending on the type of data relevant for the development of a knowledge-based system. According to Wielinga *et al.* (1987) different types of data can be identified. Among these types, verbal data is frequently used as language is the most natural way of expressing knowledge.

Besides obtaining data from experts, such tools should provide them with the autonomy necessary to enable them to express their knowledge. A natural method for the expert to express her/his knowledge may vary with the expert's culture and character, and the types of knowledge being elicited. Also, personal problems may occur such as experts who are not able to verbalise their knowledge or sometimes are not willing to show how they reason when solving their job problems.

A large number of knowledge elicitation techniques can be found in the literature, but they are not interchangeable. Different types of data may require different elicitation techniques and on the other hand, different techniques may produce different types of data as the result of the elicitation process. Some techniques, such as interviews, which are very efficient in many domains may fail in situations where the knowledge is difficult to verbalise. In those situations other techniques may be applied such as construct elicitation, card sorting and laddering. Although the purpose of this review is not to study knowledge elicitation techniques in detail, the main techniques will be briefly explained.

2.3.1 Interviewing techniques

Interviewing techniques are frequently used to elicit verbal data. An interview, though, is not so easy to perform. It requires a lot of skills on the part of both the interviewer (the knowledge engineer) and the interviewee (the expert). Different types of interview are described in the literature. The most common types mentioned are: structured, semi-structured and unstructured. Although the authors agree on these three types of interview the definition of each interview type may vary from one author to another. Following are some of these definitions.

2.3.1.1 Structured interview

Structured Interviews are strict interviews which demand from both, the knowledge engineer and the expert, a high level of knowledge about the domain area. They are interviews in which the interviewer asks the same questions in the same words and in the same order for each interviewee (Cordingley, 1989). Another definition says that in a structured interview the knowledge engineer tries to elicit all the knowledge relating to some concept or model by continuously asking for clarification, explanation, consequences, justifications, instances and, equally, counter instances (Wielinga *et al.*, 1987).

Welbank (1990) mentions that in a structured interview, specific, more closed questions are asked, not necessarily in a rigid order but according to a structured plan. According to the definitions, the goal of a structured interview is to ask the expert specific questions which the knowledge engineer should prepare before the interview which would require that the knowledge engineer had a good understanding of the domain knowledge.

Because the order of the questions is also an important factor the knowledge engineer should be aware of the fact that it is not only important to ask the right questions but also to choose the right moment. The expert could also feel the pressure of a structured interview since s/he would have to answer strict questions and not be allowed to change the subject to another topic. A structured interview would be appropriate when the knowledge engineer had a good view of the domain and the questions to ask the expert.

2.3.1.2 Semi-structured interview

A semi-structured interview does not demand the knowledge engineer to conduct the interview so rigorously. Cordingley (1989) defined semi-structured interviews as those where there is a list of questions to be asked, but the order in which they are covered and the words used to express them may vary from interview to interview. In this view a semi-structured interview is more flexible than the structured interview since the order in which the questions are asked is not relevant. This would give the expert more independence to talk about the domain, which could be an advantage or a disadvantage depending on the domain. Since the knowledge engineer would not prepare the order of the questions beforehand, s/he would need to figure out when would be the right time to ask specific questions.

2.3.1.3 Unstructured interviews

The unstructured interview gives the expert more control of the interview making the knowledge engineer more of a listener rather than a interviewer. Cordingley (1989) mentioned that unstructured interviews are designed to allow interviewees to cover topics in largely their own way. They provide "capacity for surprise" for some interviewers who have an idea of the kind of information that is needed and are prepared with a set of seed questions, prompts and probes. The idea of seed questions to direct the interview is also mentioned by Welbank

(1990) who affirmed that, in an unstructured interview, the knowledge engineer asks very general questions and allows the expert to answer them in whatever way feels most comfortable, asking supplementary questions as necessary to draw out the conversation.

An unstructured interview could be used when the knowledge engineer wants to know generally about the domain and does not want or does not have enough knowledge of the domain to ask specific questions. Although seemingly easy to perform, the unstructured interview requires from the knowledge engineer the ability to bring the expert back to the main topics when the interview is getting out of his/her control.

Comparing the three interviewing techniques mentioned above, although the definitions show slight differences, it seems that the structured interview is the most "rigid" interview and is used when the interviewer can anticipate the topics to be put to the interviewee. The semi-structured interview is more flexible, but it still requires anticipated topics from the interviewer. The unstructured interview is not so strict and allows the interviewee to express her/his knowledge in a more independent manner. In the same way that knowledge elicitation techniques depend on the type of data to be elicited, different types of interview may be required for different situations. Also, different types of interviews may be combined to satisfy specific requirements. In fact, the combination of different interview techniques may provide a better result than the use of only one technique.

2.3.2 Protocol analysis

Protocol Analysis is often considered in the literature as an elicitation technique, but there are no formal definitions or explanations of how to use this technique. Its definition may assume several forms: the standard description of a task; the task which is described; the performance of a task, what the experts says while performing a task etc. In Garg-Janardan & Salvendy (1987), protocol analysis is defined as the collection of information from subjects by having them "think aloud" or introspect and verbalize. In Welbank (1990) the idea of having the experts thinking and verbalizing their thoughts is also mentioned. Johnson *et al.* (1987) say that, historically, protocol analysis derives from the attempt by psychologists at the turn of the century to use introspective methods in order to gain an understanding of the mental process.

These definitions are similar as they emphasise the link between cognitive processes and verbalisation, or the connection between the expert's thoughts and the descriptions of his/her actions. According to Ericsson & Simon (1984), there are three types of verbal protocols. The first is described as "Think-Aloud" protocol in which some questions or problems are proposed to the person being interviewed and s/he is asked to think aloud while working on the problem given. The second is called "Talk-Aloud" protocol in which the persons being interviewed are asked to say out loud whatever they are saying silently to themselves. The last type of protocol is "Retrospective Reporting" which is done after the activity is finished. The authors recommend the use of think-aloud and retrospective reports simultaneously as they are expected to contain basically the same information. It seems that these three types of protocol also agree on the fact that protocol analysis is the connection between what the experts **say** about their thoughts and actions and the **actual** thoughts and actions. The main difference between them is when the protocol analysis is performed, either during or after the expert's task.

The three types of protocol analysis mentioned above may be compared to the interviewing techniques in the sense that both produce a verbal protocol as a result. In the case of an interviewing session the result is more likely to be called a transcript rather than a verbal report in order to avoid misunderstandings between the two techniques. Although producing similar results, the two techniques can be differentiated by the method they use to elicit knowledge. Protocol analysis is more concerned with the way people verbalise their thoughts and uses theories such as Behavioursim and Rationalism to base the elicitation process (Ericsson & Simon, 1984). Interviewing techniques may be more directly applied, not demanding the same level of understanding about the psychological behaviour of the expert.

In Garg-Janardan & Salvendy (1987), some pitfalls of protocol analysis were mentioned. Timing of verbalizations, directedness and content of verbalizations and amount of intermediate processing required were mentioned as factors that can affect the consistency and completeness of verbal reports. But the authors also affirm that despite its weaknesses protocol analysis has been successfully used to elicit knowledge for several expert systems.

2.3.3 Construct elicitation

The term "personal construct" originates from George Kelly's work (Kelly, 1955), The Personal Construct Theory. Based on psychological evidence, the theory of Personal Constructs indicates that people seek to predict and control events by formulating theories, testing hypothesis, and weighing evidence. Kelly asserted that an individual's perception of the world, other events, individuals and situations was represented in the form of constructs.

According to the analysis of Kelly's work in Cordingley (1989), Kelly developed a technique for exploring the way people make sense of the world called Repertory Grid Technique. Later, Kelly developed the Theory of Personal Constructs consisting of fundamental postulates and eleven associated corollaries. A construct is a mental "tool" which allows a person to discriminate between elements of his/her world. An element is a personal observation or experience of the world. For example, in the Computer Science world, a computer program would be an element and "hard to use/easy to use" would be a construct.

A repertory grid as presented by Kelly would be a cross reference between elements and constructs. The first step would be to choose elements relevant to the domain. The expert should then divide the elements into groups of three according to their type and level of complexity. For each group of three the expert should name the construct in which the elements differ from each other. This should be done for each group until the grid is complete. Some knowledge acquisition tools use Kelly's technique in order to acquire knowledge from the expert as, for example, the Expertise Transfer System described in Boose (1985). The goal of this system is to elicit the knowledge from the expert, analyze it and construct production rule knowledge bases. The idea here is not to analyze the performance of such systems but to mention the use of the repertory grid as the knowledge elicitation technique used in these systems. Although using the repertory grid as the basis for the knowledge elicitation process in the Expertise Transfer System, Boose mentioned that there are some limitations in the repertory grid such as the difficulty in verifying that a sufficient set of constructs has been elicited. Inappropriate constructs are relatively easy to eliminate from the system, but errors of omission are hard to detect. Also, because constructs are personal, a grid done by one expert may differ from a grid about the same domain done by another expert. This fact could be seen as an advantage of the repertory grid since it could bring up points which were omitted by either expert.

It should be said that the repertory grid technique is not always suitable for the knowledge elicitation process. It would depend on the domain characteristics as well as on the knowledge engineer's understanding of the domain in order to evaluate the elements and the constructs. In domains where the elements could be easily separated, the repertory grid could be a possible choice for a knowledge elicitation technique. Another way of using the repertory grid would be as a tool for evaluating the knowledge elicited from the expert in order to check if all the concepts were extracted.

2.3.4 Card sorting

Card sorting is another technique which uses elements of a domain to understand how the expert conceives his/her world. The elements are either elicited from the expert or taken from an analysis of the domain and are then written on small cards. Card sorting can be done in several ways. The cards can be divided into two piles and the piles labelled according to the main characteristics of their elements. The piles are then subdivided until the process can go no further. Another way is to give the expert two cards and ask whether the elements are the same or different. If they are the same they are put in the same pile, if different in another one. A new element is chosen and the session proceeds as before until all the elements are used. The last way is to give the expert all the elements to sort into as many relevant piles as possible. The piles should also be labelled according to the main characteristics of their elements. According to Welbank (1990) this technique is good for eliciting the conceptual structure of the domain, but has been considered rather artificial by experts (Freeman, 1985).

In the same way as construct elicitation this technique is not suitable for all the domains. It might be more useful as a prompting device to check the knowledge obtained.

2.3.5 Laddering

Laddering is based on the process of generating hierarchies of concepts. The hierarchy is constructed according to the questions put to the expert by the knowledge engineer. These questions generate subordinate and superordinate concepts as well as concepts at the same level in the hierarchy. To get superordinate concepts the expert is asked of constructs "Why ... ?"; to get subordinate concepts the expert is asked "How ... ?"; and to get concepts of the same level the expert is asked for "Alternative examples of ...".

The questions start with a seed item for which the expert expresses preference for one of two poles and then the knowledge engineer asks why s/he prefers one side to the other side. The questions continue until the knowledge engineer gets to core constructs i.e., ones for which the expert can answer no more "why" questions but reiterates the same answer to succeeding "why" questions. The knowledge engineer should not question the expert's fundamental beliefs, such as religion.

Again this technique may not be appropriate for all kinds of domain. It might in fact generate some problems if the knowledge engineer did not know how far the questions should go. In Burton (1988), laddering is characterised as a goal decomposition technique. But according to Cordingley (1989) the technique would be appropriate when any kind of hierarchy of concepts is to be created in the domain, not just for decomposition of goals.

2.4 Analysis techniques

Although considered interchangeably in the literature, elicitation and analysis techniques have particular characteristics. As mentioned in the latter section, elicitation techniques are used to obtain the data from the expert. The result of this process is the unprocessed knowledge which has to be studied and understood before being represented in the knowledge base. Analysis techniques are used for the purpose of evaluating this unprocessed knowledge. Another characteristic differentiating the two techniques is the fact that the presence of the expert is not necessarily required to perform the analysis process.

In Wielinga *et al.* (1987), the authors mention the fact that the data have to be interpreted in order to make them suitable for current knowledge representation techniques. Two methods for performing the data interpretation are given: data-driven and model-driven. The data-driven method proceeds through a bottom-up analysis of data to find a structure that can support the problem-solving process. The model-driven method starts with a more general model of the problem solving process and attempts to find the relevant knowledge elements in the data.

An important data-driven analysis activity is concerned with structuring the domain specific concepts. Structures are built using general relations such as ISA or Consists-of. Such structures of domain concepts reflect the static structure of the domain knowledge. The following techniques are examples of how these concepts and relations can be obtained.

2.4.1 Discourse analysis

One technique which seems to fulfil the requirements of the activities mentioned above is discourse analysis. Because there is no specific definition of what the technique really is, i.e. an elicitation technique, an analysis technique or a tool for evaluating verbal data; discourse analysis will be explained through some examples of its use rather than by a formal definition.

In Kuipers & Kassirer (1987), a verbatim transcript was used as the raw data of a medical experiment. As it was transcribed, the transcript was broken into short lines that corresponded roughly to meaningful phrases in the explanation. Selections were made of excerpts in which the subject seemed to be concentrating on the explanation and presenting medical knowledge rather than expressing the expert's opinion about his/her own mental processes. The analysis of an excerpt takes place in two stages. Firstly it identifies the objects and relations in the domain that the subject is referring to; and secondly it identifies the causal relationships that are described in the segment.

In the second example which is based on the work of Belkin *et al.* (1987), discourse analysis was also used as a method to collect data from transcripts. The transcripts were the result of taped interviews. The analysis was done by dividing the transcripts into utterances, which they called units of analysis. An utterance can be defined as a speech sequence by one participant during the conversation. The groups of utterances were highlighted according to concepts or themes that were the focus of attention in the transcript. The purpose of the analysis of each interview was to identify the tasks and functions being carried out at each level and the knowledge required to perform the tasks and functions as well as the interactions between functions and relations. The authors justified the use of discourse analysis mentioning the need to analyze the recorded observations at the discourse level.

The approaches mentioned are examples of how a transcript may be transformed into a more detailed representation of the domain expertise. To do so, both mentioned the selection of statements made by the experts about important concepts of the domain. The relationships between such concepts and the functions performed by the experts were mentioned as part of the analysis. Although the examples did not establish a formal way of performing the discourse analysis, they show that the technique is based significantly on the selection, from the transcripts, of primitive elements such as concepts, relations and functions according to the emphasis given by the experts to these elements.

2.4.2 Semantic networks

In his historical review, Brachman (1979) gives several examples of how semantic networks have been used and defined since their first appearance in Quillian's (1966) work. Quillian was one of the first researchers to suggest that human memory could be modelled by a network of nodes, and proposed a processing model for memory retrieval.

A definition by Jackson (1990) is that a semantic network is a network used to structure general kinds of information. These are collections of nodes and links in which the nodes stand for concepts and the links stand for relationships between them. An example of a semantic network is shown in figure 2.1.

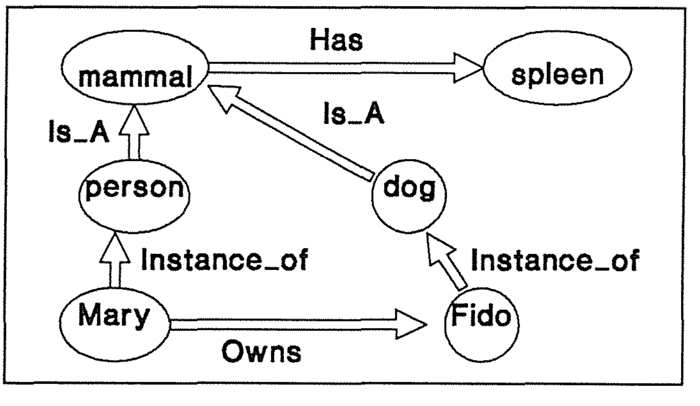


Figure 2.1. An Is_A semantic network.

The author mentions that the semantic networks are so named because they were originally employed to represent the meaning of natural language expressions. This facility of transforming natural language into a knowledge representation adds a benefit to the use of semantic nets. Other advantages are the flexibility to add, modify and delete new nodes and arcs where appropriate, and the ability to reason and make assertions about one node and its relationship with another node where no direct arc exists between the two nodes. According to Brachman (1979) the different types of semantic net primitives (concepts and relationships) can be considered to form a set of levels, or viewpoints, according to the kind of knowledge the nets represent (figure 2.2).

Level	Primitives
• Implementational	- Atoms, Pointers
• Logical	- Propositions, Predicates, Logial Operators
• Epistemological	- Concept Types, Conceptual Subpieces, Inheritance and Structuring Relations
• Conceptual	- Semantic or Conceptual Relations, Primitive Objects and Actions
• Linguistic	- Arbitrary Concepts, Words and Expressions

Figure 2.2. Levels of semantic networks (from Brachman, 1979).

In implementation level networks, links are merely pointers, and nodes are simply destinations for links. This level assumes the network to be only a data structure, making no important claims about the knowledge structure. In the logical level nets, nodes represent predicates and propositions. Links represent logical relationships as, for instance AND. In the epistemological and conceptual level nets, links represent semantic or conceptual relationships. Those are claimed to be the "real" semantic nets whose primitives are word-senses and case relations. The primitive elements in the linguistic level are language-specific.

According to this classification, semantic nets may be used for several purposes according to the kind of knowledge which is being analyzed. In terms of knowledge representation this flexibility has been identified as one of the disadvantages of semantic nets. The lack of formalism of the technique would not fulfil the needs of a formal knowledge representation. However, semantic nets seem to be a very useful technique to represent the primitives of a domain and could be used as an input for a more formal representation such as production rules or frames. Also, semantic nets and discourse analysis could be used as techniques to refine the data obtained from the expert during the elicitation process.

2.5 KADS methodology

KADS has been frequently mentioned in the literature as a very promising methodology for building knowledge-based systems. In Hickman (1989a), they take the view of a knowledge-based system development in KADS as a modelling activity. The author mentions that a knowledge-based system should model not only the expertise but also other real world systems such as the organisation, the user and the user/system interaction, in order to obtain a commercially viable system.

It is common to find references which give more attention to some different components of the KADS methodology. Among these components, the most emphasised are the model of expertise and the library of interpretation models.

2.5.1 The models in the KADS methodology

Because of the model-driven approach adopted by KADS, the methodology is explained as a set of different models. The methodology is often visualized as a constructive activity containing several components. Figure 2.3 shows the sequence of the models which comprise the KADS methodology.

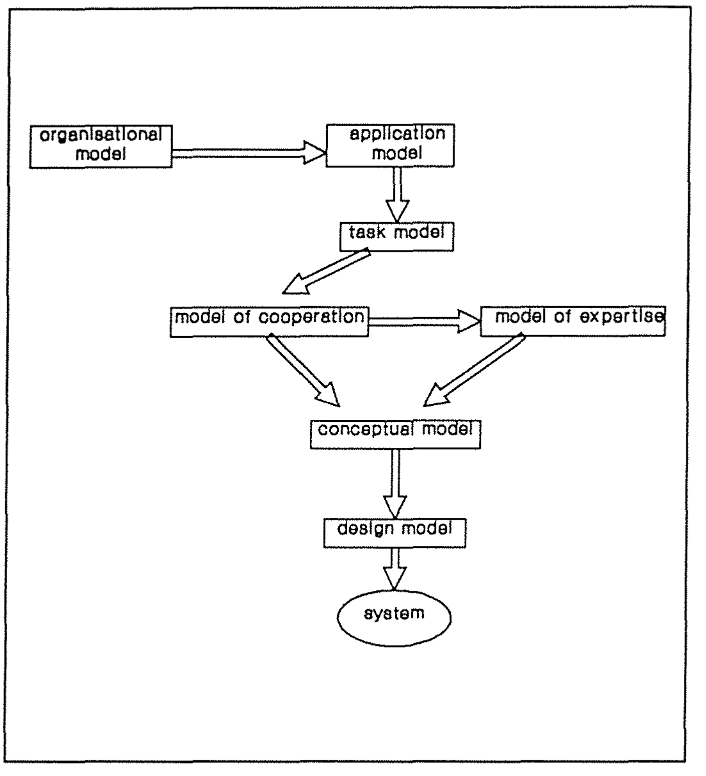


Figure 2.3: Intermediate models in the KADS methodology (from Wielinga, 1991).

Following is a brief explanation of each of these models, according to Wielinga et al. (1991):

- **Organisational Model**: This model provides an analysis of the socio-organisational environment which will be analyzed. It consists of a description of the functions, tasks and bottlenecks in the organisation.

- **Application Model**: This model defines what problem the knowledge-based system should solve in the organisation and which is the system goal.

- **Task Model**: The task model specifies how the goal of the system is achieved by a number of tasks that the system will perform.

- **Model of Cooperation**: This model consists of a more detailed description of the tasks acquired in the task model.

- **Model of Expertise**: This model is based on two premises. The first affirms that is possible to distinguish between several generic types of knowledge; and the second, that these types can be organised into layers. As the model of expertise is one of the main components of the KADS methodology, it will be explained in more detail in section 2.5.2.

- **Conceptual Model** : The conceptual model, as it is shown in figure 2, is the result of merging the model of cooperation and the model of expertise. Conceptual models are abstract descriptions of the objects and operations that a system should know about.

- **Design Model**: The design model consists of the computational and representational techniques that the system should use to achieve a specific behaviour. In building a design model it is necessary to take into account some requirements such as speed, hardware and software.

Although very popular, some weaknesses in the KADS methodology could be mentioned. For example, the user model is not clearly specified. One could say that the user model may be part of the organisational model, but this is not very well explained. Another point is the difficulty in understanding how some of the models are constructed. The meaning of the conceptual model, for instance, is not obvious due to its high level of abstraction. However, these weaknesses seem to be less important than the advantages of using such methodology.

In practice KADS has been used with success, as mentioned in Hayball & Barlow (1990), where the models of the methodology assisted in the conceptualization and construction of a knowledge-based system. Among these models two of the more important are the model of expertise and the library of problem-solving activities also called library of interpretation models. These two areas are the subject of the next sections.

2.5.2 The model of expertise

The major question in a knowledge-based system is whether or not the expertise has been acquired and analyzed in a proper way. From this perspective, several methodologies which claim to fulfil the needs of a knowledge-based system development are very similar to the traditional methodologies (Guida & Tasso, 1989; Basden, A. 1989). One of the problems seems to be that these methodologies take a traditional approach and try to adjust the needs of a knowledge-based system to it. Because these systems require a much more detailed interface with the user and also add the expert's role to system development, it is necessary to have a different way of representing the knowledge. In this sense, the modelling of expertise knowledge is a crucial part in a knowledge-based system development.

In the KADS methodology the process of modelling the expertise knowledge starts with the distinction of different types of knowledge. This distinction indicates different ways in which the knowledge can be viewed and used. According to Wielinga *et al.* (1991), it is possible to distinguish between several generic types of knowledge according to different roles that knowledge can play in the reasoning processes. The first distinction made is between static and control knowledge.

Static knowledge describes a declarative theory of the application domain. The control knowledge comprises the different types of inferences that can be made in this declarative theory (first type of control knowledge) and knowledge representing elementary tasks (second type of control knowledge). The third type of control knowledge is called strategic knowledge and comprises the problem-solving processes. This division of knowledge into different levels takes into account the fact that knowledge may assume different forms according to its characteristics. The first category of knowledge is generally regarded as the static level for containing the concepts, relations and attributes of the domain expertise. This level should not contain actions or problem-solving strategies and is implementation independent. The second level is considered as the dynamic part of knowledge for containing the inferences that can be made about the concepts and relations of the static level. The tasks the system will perform are also part of this dynamic knowledge and describe the sequence of actions taken to achieve a specific goal. The problem-solving methods or strategies applied to achieve the goals complete the dynamic knowledge.

It should be said that the word "task" is being used here as a way of expressing the actions that the system has to perform to achieve a desired purpose. The fact that the word task has been used for expressing different concepts can lead to some misunderstandings. In the context of the KADS methodology some examples of the use of the word task can be mentioned as: task model, task layer, task structure and library of generic tasks. The task model as explained in section 2.5.1 is one of the models which constitute the KADS methodology. The task layer is one of the layers of the model of expertise and the task structure is a hierarchy of tasks and subtasks and is part of the task layer. And the last use of the word task is for the library of generic tasks which is a set of activities which a system could perform in order to achieve a specific goal. The intention here is to use the word task in its plain meaning and when some other sense is required the concept should be explained and its meaning clarified.

Another classification of knowledge was done by Brachman (1979) and was mentioned in section 2.4.2 (Semantic Nets). The author classified the levels of semantic networks according to the kind of knowledge they represent. The same classification was done by Hayward (Hayward *et al.* 1987). The author mentions the classification of different kinds of knowledge as a guide to the construction of the model of expertise. Following is the classification according to Hayward's view:

- Knowledge Identification: This level corresponds to recording what one or more experts report on their knowledge. Different experts may be reported in a different way.
- Knowledge Conceptualization: It is the formalization of the knowledge in terms of conceptual relations, primitive concepts and conceptual models.
- Epistemological Analysis: This level comprises epistemological primitives representing types of concepts, types of knowledge sources, structuring relations(such as hierarchical relations, inheritance), and types of strategy.
- Logical Analysis: This level expresses the knowledge which is

responsible for inference making.

- Implementational Analysis: This level describes mechanisms of knowledge implementation.

These five levels of analysis aim to produce a framework of the expertise. This framework is used to separate the different kinds of knowledge into the four layers of the model of expertise (figure 2.4).

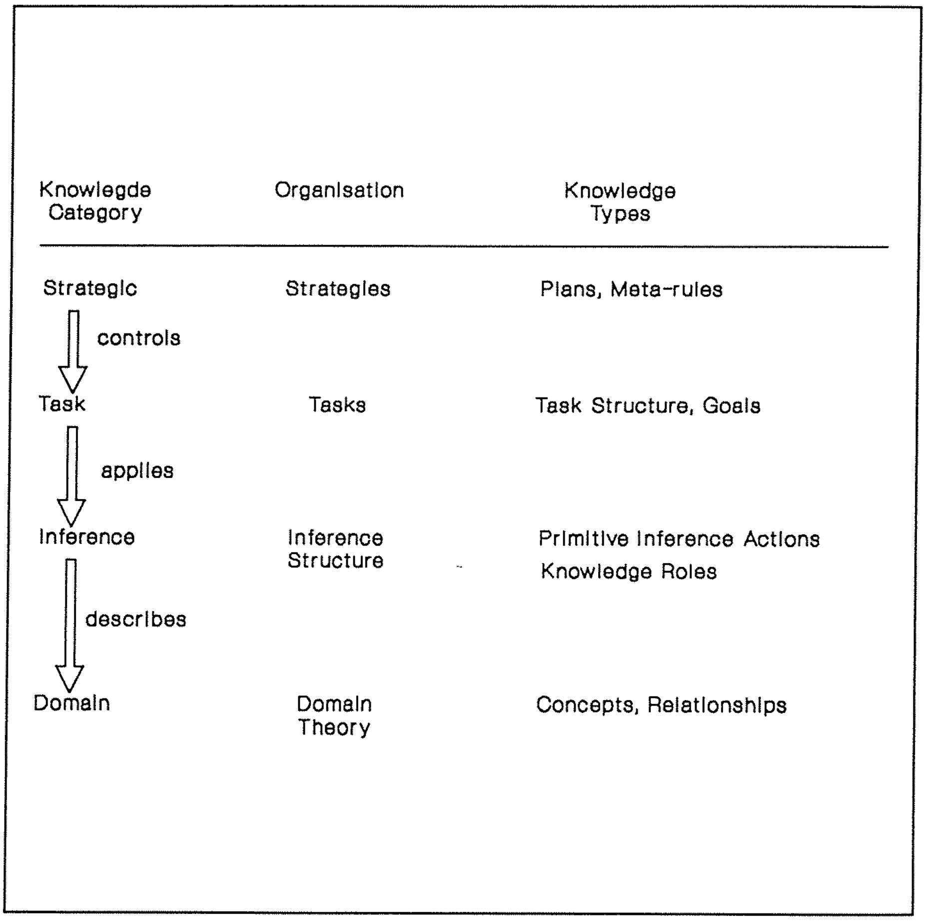


Figure 2.4. The KADS four-layer model (in Wielinga, 1991).

- **The Domain Layer** : This is the lowest layer of the model of expertise and represents the conceptual level of a domain theory. This domain theory is described by a set of **primitives** which are: concepts, properties, and relations. The most frequent relations are the ISA and part_of relations but others may be relevant such as, consists_of, depends_on etc. The knowledge in this layer is task-neutral in the sense that there are no means to control its use. The domain layer should contain all the generic facts on which the reasoning is based.

- **The Inference Layer:** This layer contains meta-knowledge about the domain layer. It can be viewed as a set of inference making functions such as abstraction, specification etc. These functions are called primitive inference actions. These functions have input and output arguments which are called knowledge roles and are usually named according to their role in the problem solving process.

By combining primitive inference actions and knowledge roles an inference structure can be constructed that reflects the required inference competence to accomplish a problem solving task. An inference structure describes which inferences can be made but not how or when they are made. Figure 2.5 shows a simple inference structure for a hypothesis generation and verification task.

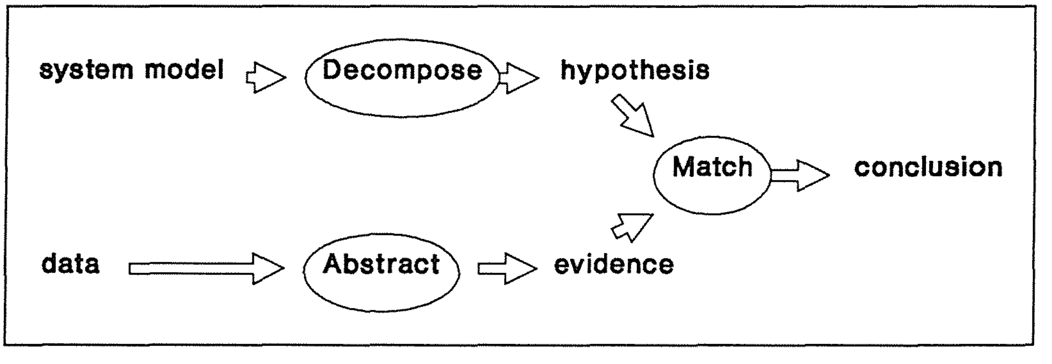


Figure 2.5. Inference structure for hypothesis and testing (in Breuker and Wielinga, 1989).

- **The Task Layer:** This layer contains the specification of how the inferences can be used in the problem solving process. According to Wielinga et al (1991) the prime type of knowledge in this layer is the task. Tasks can achieve a particular goal. Tasks represent fixed strategies for achieving problem solving goals. Here task is understood by a composite problem solving action. This implies its decomposition into sub-tasks leading to a task structure which is a set of sub-tasks.

- The Strategic Layer: This layer determines what goals are relevant to solve a particular problem. The strategic layer should suggest new lines of approach or attempt to introduce new information in case an impasse is reached in the problem-solving process. It involves the dynamic planning of task execution.

2.5.3 The library of problem-solving activities:
- The interpretation models -

According to Wielinga *et al.* (1991), interpretation models are models of expertise with an empty domain layer. They are templates which contain specifications of the inference and task layers for a particular task. Here the use of the word task appears again. In this review the use of the term task in the context of the library of the interpretation models will be replaced by the term problem-solving activities to avoid misunderstandings. As a top-down knowledge acquisition tool, the interpretation models can guide the interpretation of verbal data obtained from the expert.

The major division in the interpretation models is between analysis and synthesis problem-solving activities as shown in figure 2.6. Analysis tasks have the goal of establishing unknown properties or behaviour of an object within the domain. Analysis of an unknown property may be diagnosis of a fault. Analysis of behaviour may be the prediction of a future state. Synthesis problem-solving activities aim to compose a structural description as a possible object within the domain. Typical synthesis activities are design and planning. Each of these problem-solving activities have a template which represents the problem-solving process for such activities. Although the library is a very helpful tool for constructing the model of expertise, the process of choosing such activities involves some difficulties. It is not an easy job to decide which activity is the most suitable for a particular domain. Sometimes the complexity of the domain may require the choice of several problem-solving activities which could result in a very complicated template. Another point is the fact that not all problem-solving activities are well developed in the KADS library. One of the most developed activities is the diagnosis activity. But, on the other hand, activities such as design are poorly developed.

```
is_a_problem_solving_activity
system analysis
|       identify
|       |       classify
|       |       |       simple classify
|       |       |       diagnosis
|       |       |       |       single fault diagnosis
|       heuristic classification
|       systematic diagnosis
|       causal tracing
|       |       localisation
|       |       |       |       multiple fault diagnosis
|       |       |       assessment
|       |       monitor
|       predict
|       |       prediction of behaviour
|       |       prediction of values
system synthesis
|       design
|       |       hierarchical design
|       |       configuration
|       planning
|       modelling
```

Figure 2.6. The KADS library of problem-solving activities (in Breuker *et al.,* 1987).

This does not mean that the library is not helpful but it has some weaknesses as considered in Wielinga *et al.* (1991). The importance of the generic problem-solving activities as a guide to the modelling of knowledge is mentioned by other authors (Clancey 1985; Stefik *et al.*, 1983; Mattos 1991). In this research work the problem-solving activities are considered as a guide to the construction of the model of expertise.

2.6 Other methodologies for knowledge-based systems

Although having several supporting tools the development of knowledge-based systems has not yet reached a widely accepted building methodology. It is possible to identify in the literature some of the branches of such methodologies.

In Kahn & Bauer (1989) and Bratko (1989) the Rapid Prototyping paradigm is presented. As described by Roberts (1990), in rapid prototyping the knowledge engineer elicits knowledge from the expert and builds it directly into the system. The expert then tries the system and points out any faults or omissions which are corrected. This process is repeated until the prototype reaches the desired state.

One of the problems of rapid prototyping is that the process is not a methodology but a tool for demonstrating to the user and expert what they can expect from the system. This fact can lead to advantages as well as disadvantages. It is useful to make sure that the developers are on the right track and to check with the experts if the problem-solving process is working properly. But, on the other hand this could give wrong ideas about the performance of the system which could cause a great disappointment in the end.

Another disadvantage is that when using rapid prototyping the analysis phase is usually done very quickly because of the hurry in constructing the prototype. This could lead to a poor knowledge base and inadequate problem-solving processes.

According to Hickman (1988), rapid prototyping is an *ad-hoc* development which often leads to poorly specified, poorly designed systems, with low portability and a high maintenance cost. The point here is not to exclude prototyping from the knowledge-based system lifecycle, but to use the prototype as a tool and not as the methodology. A study of the use of the model of expertise and prototyping in the knowledge-based system lifecycle can be found in Kemp & Kemp (1991). Another discussion of the main advantages of prototyping as well as its pitfalls is presented in the study of Carey & Currey (1989).

Other methodologies try to adapt the traditional concept of the software life cycle to knowledge-based systems development. According to Guida and Tasso (1989), the knowledge-based system technology presents some key concepts which make it different from the traditional software as, for instance, the difficulty in identifying the fundamental systems components (the modules, in traditional systems) or the lack of concrete experience with the design and development of real-size projects. This seems to be the reason for several approaches that adopt a different vision in the development of knowledge-based systems.

In Bylander & Chandrasekaran (1987), the Generic Task Approach was adopted as a way to organise and represent the knowledge needed to solve certain kinds of problems. The main idea behind this approach is the fact that different tasks (such as Diagnosis and Design) require different representations. In Musen *et al.* (1987), the Method-to-Task Approach claims that it is possible to create a systematic domain terminology and a general model of the tasks of the domain, i.e. a task-based conceptual model.

Comparing the approaches mentioned to the KADS approach, it is possible to notice some similarities in terms of the importance of the generic tasks or, as called in the KADS methodology, the library of problem-solving activities. However, according to Karbach *et al.* (1990), there are major differences between the modelling approaches. KADS would be the most flexible one due to its lack of operational semantics. Also, in KADS, the conceptual model is supposed to model the expertise rather than the system. According to the author the combination of the approaches would be a very promising methodology.

2.7 Conclusion

There is still no widely used methodology for building knowledge-based systems. However some tools and techniques have been proposed in order to fulfil the requirements of such systems. Although not complete, the library of interpretation models in KADS promises to be a very helpful tool to guide the construction of the model of expertise. The division of the knowledge into different types in the model of expertise was a decisive point in the sense that the domain knowledge needs a categorisation according to the role that it plays in the reasoning process.

In terms of knowledge acquisition the intention was to mention that a formal definition does not exist and different authors may consider knowledge acquisition differently. This is also true of the concept of knowledge elicitation.

Here, *knowledge acquisition is considered as a constructive process of collecting (eliciting), analyzing, interpreting and modelling the data.* The result of the elicitation process is the unprocessed data obtained from the expert through elicitation techniques such as interviews. The analysis concentrates on the evaluation of the material obtained in the elicitation phase, in order to refine the knowledge obtained. The interpreting and modelling processes aim at representing the domain elements obtained during the analysis process.

Chapter 3: Domain problem

3.1 Introduction

It was suggested that a knowledge-base system could be built to assist farmers with summer-autumn management on a dairy farm. The system would concentrate on this period during which dairy farmers are faced with many crucial decisions. Among these decisions, one of the most important is the drying off date, that is when farmers stop milking the cows.

In order to understand the consequences these decisions may have for milk production, a typical dairy system in New Zealand is described. The structure of the New Zealand Dairy Industry is also explained to illustrate its importance to the New Zealand economy and the decisive role played by dairy farmers. Finally, a general description of the phases of the dairy project is presented.

3.2 Dairy industry overview

The New Zealand Dairy Industry is a farmer owned cooperative which is managed by the New Zealand Dairy Board, an organisation which aims to maximise the profit of its farmer shareholders (Bull, 1992; Hall, 1991; and Weeds, 1993). The industry is vertically integrated from farmers through processing companies to distributors (Figure 3.1). Approximately 13,800 dairy farmers supply raw milk to 16 dairy companies where the milk is transformed into processed products. The companies provide processed products to the New Zealand Dairy Board which is responsible for the marketing, distribution and sale of the products.

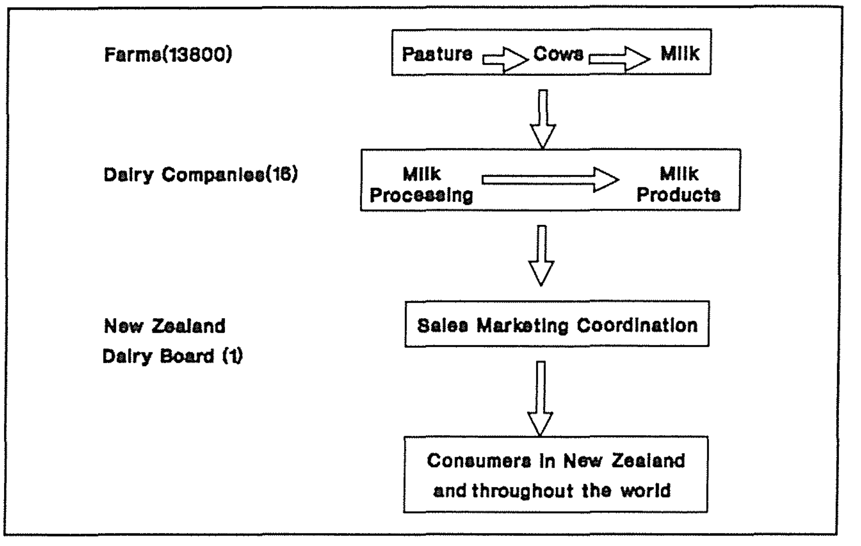


Figure 3.1. The vertical structure of the New Zealand Dairy Industry (Adapted from Holmes, 1993).

The New Zealand Dairy Board exports about 90% of New Zealand's total milk production, which corresponds to 24% of the total world's trade in dairy products. Dairy products account for 25% of New Zealand trade, which represented an earning of approximately $3.9 billion and 5.3% of New Zealand's gross domestic product in 1991 (McEldowney, 1992).

Despite the importance of the dairy industry to the New Zealand economy, milk production does not receive government subsidies. This differs from the majority of New Zealand's competitors in the international trade who subsidise their local dairy farmers. This unsubsidised system presents New Zealand farmers with the advantage of not depending on export incentives which might be eliminated during an economic crisis. On the other hand, this has the disadvantage of letting New Zealand farmers be fully exposed to price variations of the world market. Despite this, the New Zealand dairy industry has operated on a fairly stable basis and has shown an expansion since 1988/1989 due to increasing product prices (Table 3.1).

Table 3.1: Summary of statistics for the New Zealand Dairy Industry (adapted from Dairy Farming Annual: 1987/88/89/90/91/92)

Year	86-87	87-88	88-89	89-90	90-91	91-92*
- Cows Milked (Thousands)	2,280	2,236	2,138	2,141	2,225	2,250
- Cows/ Herd	151	153	156	160	166	-
- Milk Production (litres/cow)	2,928	3,207	3,004	3,220	3,190	3,313
- Milkfat Production (kg/cow)	138	154	143	152	152	161
Total Production:						
- Milk Production (million litres)	7,031	6,921	6,533	6,883	7,078	7,455
- Milkfat Production (thousand tonnes)	332	333	311	330	342	363

* data estimated on 1 May 1992.

3.3 Dairy production from pasture - An overview of New Zealand dairy systems

Dairy production in New Zealand is based on the conversion of pasture into milk by cows grazed *in situ*. The quantity of supplementary feed per cow is small in comparison with dairy farms in many other countries where the animals are housed for part of the year and given conserved forage and concentrates. This is possible because of New Zealand's favourable climatic conditions which allow pasture to grow throughout the year (Holmes & Wilson, 1984).

Since grazed pasture provides more than 90% of all feed eaten by dairy cows (feed is complemented by silage, hay and crop), New Zealand farmers have developed a low cost production system which is based on the balance between feed supply and feed demand. The key point in this kind of system is the selection of stocking rate (number of cows per hectare) and calving date that best match the ideal balance between feed demand and feed supply (pasture and supplements).

A typical dairy farmer in New Zealand would run 169 cows on 62 hectares, producing 154 kilograms of milkfat per cow per year and 367 kilograms of milkfat per hectare per year (Farm Monitoring Report, 1992). Figure 3.2 shows the main events on a typical dairy farm throughout the season according to a standard pasture growth rate curve.

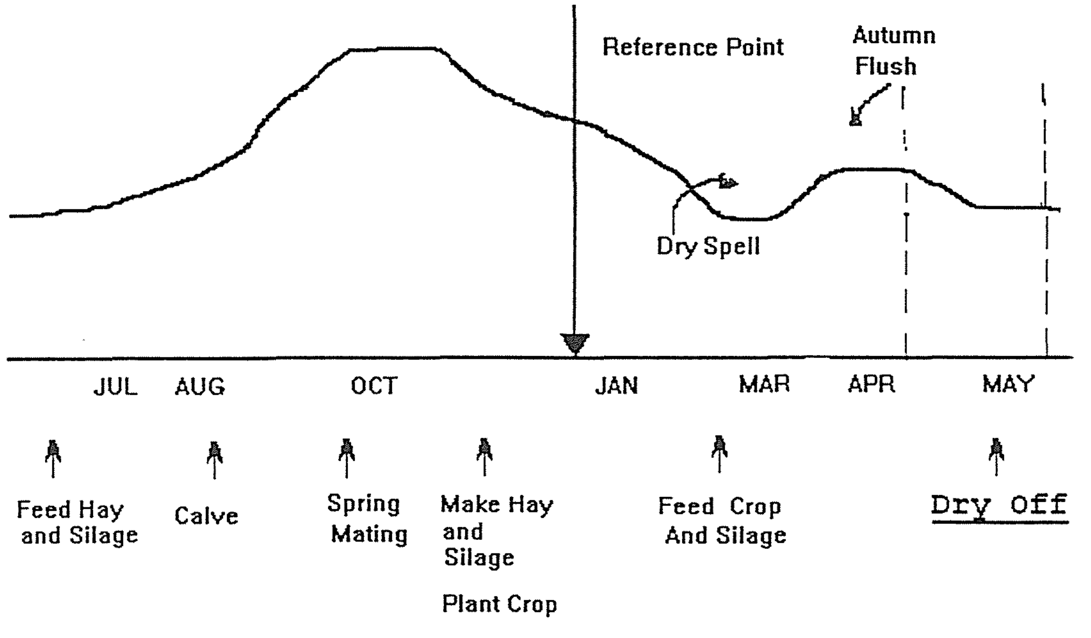


Figure 3.2. Dairy farm events in relation to the pattern of pasture growth.

The above schematic representation shows a typical dairy system in New Zealand in which all cows calve between July and September, become pregnant October to December, are milked throughout spring, summer and autumn until drying-off in Autumn between April and May. Pasture starts to be conserved in late spring in the form of hay and silage which are fed during periods of feed deficit (summer and winter). A small area of forage crop is often grown during summer to provide feed during the dry spell.

Given the high variability of pasture growth rates, farmers are faced with the problem of how best to manage events on the farm in order to make sure that cows will have enough feed during both the current and following seasons.

3.3.1 Problem domain

In terms of management purposes the dairy season represented in the diagram can be divided into three main periods. The first period is the winter management from drying-off until calving, followed by the spring management period from calving until December. The summer-autumn management period begins in January as indicated in the diagram by the reference point.

Due to highly variable pasture growth rates during this period, farmers are faced with many problems. They have to make several management decisions which include: the right time to feed supplements (hay, silage, crop), sale of unwanted cows, drying off date, body condition[1] at calving date etc. The drying off date is one of the most critical decisions farmers have to make. It determines the extent of loss in condition in late lactation (autumn) and therefore the amount of feed required during the cold period (winter) to replace that condition (Bryant, 1980).

---

[1] The physical conformation of the animal in terms of muscle and fat cover in relation to its suitability for standing the winter and the calving process.

This means that drying off the cows too late will reduce their body condition which may be difficult to replace before calving date. This would cause low production during early lactation in the current season as well as poor reproductive performance at mating time of the next season. It might take years to correct any problems. Another consequence of drying off too late is that feed that should be conserved for winter consumption (pasture and silage) would be fed earlier in order to maintain the production level.

On the other hand, drying-off the cows too early may reduce the income in the current season. Since the main goal of dairy farmers is to optimise milkfat production in the current season without risking next season's production, the drying-off date is one example of a management decision that may compromise the achievement of this goal. Making the wrong decisions during the season, specially the summer-autumn management period, may put at risk years of production ahead.

3.4 Research project: the development and evaluation of a knowledge-based system for dairy farmers

The idea of a computerized system to assist dairy farmers with the process of making management decision has arisen. The system would contain the heuristics used by expert farmers to solve the problem of how best to allocate resources through the summer-autumn period to optimise milkfat production in the current season without risking the following production. The use of knowledge-based systems technology has been seen as one of the ways dairy farmers can be helped with seasonal management (Todd *et al.*, 1993a).

3.4.1 Project goals

A knowledge-based system is being developed (Gray et al., 1992(a) and Gray et al., 1992(b)) by a multidisciplinary team from Massey University. The goal is to help a dairy farmer to make important decisions from January until the herd is dried off in late Autumn. The project has the assistance of the New Zealand Dairy Board Consultancy Officers who will be the system's users and the link between the system and the dairy farmers.

The Consultancy Officers would bring farmers more information about decision making on a dairy farm by presenting them with the system and asking them questions about the situation on their farms and what kind of problem they were trying to solve at the moment. The system would provide the farmers with an answer or advice on the actions to be taken. A prototype will initially be built to validate the content of the knowledge base.

3.4.2 Knowledge-based systems and dairy production

As described in section 3.2, the dairy season in New Zealand presents the farmers with several difficult problems to be solved due to the variability of pasture growth rates and the consequent difficulties in adapting the activities on the farm to this variability.

To evaluate the suitability of a knowledge-based system application to solve these problems the checklist provided by Beckman (1991) was used. This checklist contains the criteria used to estimate the potential for success in applying knowledge-based system technology in a given domain which are: task, payoff, customer-management, system designer, domain expert and user. The dairy farm domain, regarding the summer-autumn management period, scored highly on characteristics of the categories task, payoff and domain expert. It also scored at a satisfactory level on the categories of customer-management, system designer and user.

3.4.3 The selection of experts

The expert farmers were selected from within the Manawatu region by Dairy Board Consultancy Officers. The criteria for the selection of the farmers comprised: dairy farmers whose farms were located in different environments, had at least 10 years of experience, had achieved high production levels relative to other farmers farming on similar soil types, could articulate their management processes and could meet the time requirements associated with the project.

3.4.4 The elicitation process - semi-structured interviews -

The elicitation process was divided into three phases: the background interviews, the pilot study and a series of fortnightly interviews. In the first phase of the process, the farmers' goals, personal profiles, ownership structures, and farm resources were obtained. In the second stage, four interview sessions were held with one of the farmers with the purpose of developing a framework for the third part of the elicitation process. The aim was to obtain a general picture of the summer-autumn management.

In the last and longest phase, the farmers were interviewed on a fortnightly basis from January 1992 until their herds were dried off (May/92). These were semi-structured interviews in which each expert was asked the same questions but sometimes the emphasis varied according to the specific situation on the farm. Similarly, the order in which the questions were asked was not strict and the experts were relatively free to talk about different points in a different chronological sequence.

In the first semi-structured interview the experts were asked to explain the current state of the farm, the methods they were using to monitor the farm state, planning horizon and their plans for the future. The subsequent interviews were similarly structured, with the addition of questions to determine what farmers had done since the last interview. This provided a set of explanations about the changes on the farm, how they were being detected and what had been done to correct them. All semi-structured interviews were conducted by two participants of the group and were tape recorded and transcribed to text, resulting in a set of ten transcripts for each of the four experts.

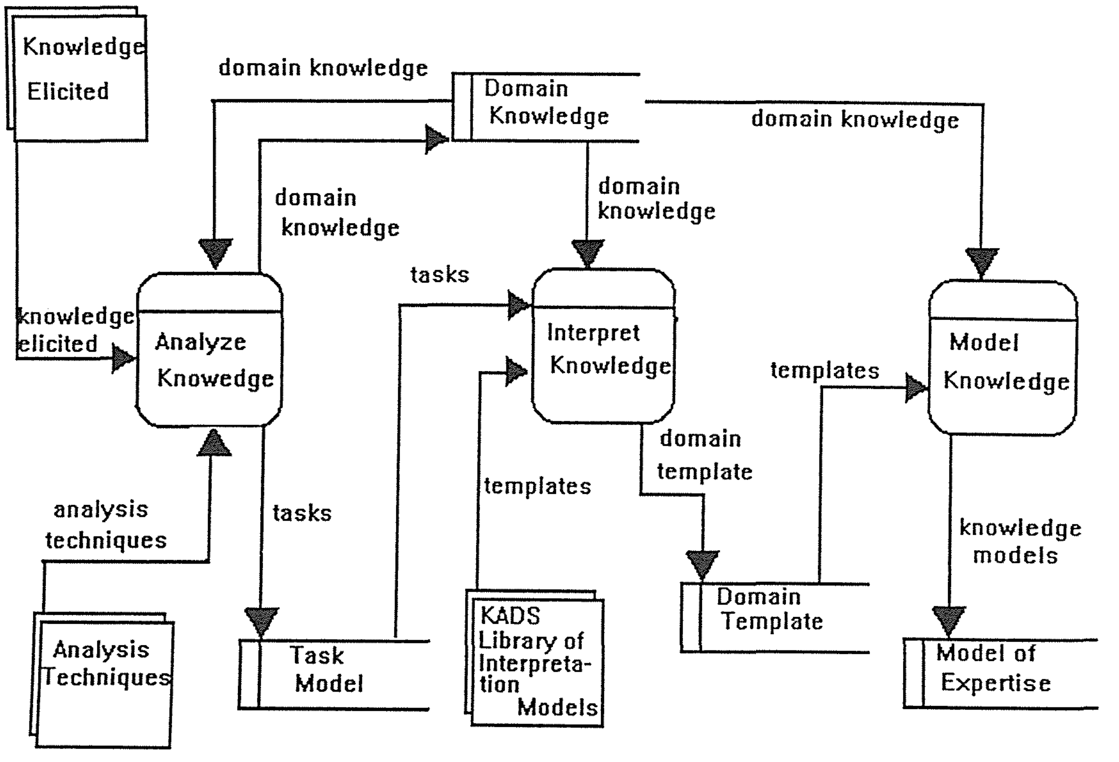
3.5 Conclusion

The Dairy Industry is a very important part of the New Zealand economy. This is mainly due to the work of the New Zealand Dairy Board which aims to help dairy farmers better achieve their production goals. The use of new techniques to help the work of the Dairy Board may have a significant impact on the economy.

Knowledge-based systems technology is one way to improve the performance of the dairy industry by helping farmers make the right decisions and consequently maximise their profits. This could have a considerable effect on New Zealand's exports, as an improvement on each farmer's production can represent larger export earnings. The use of knowledge-based systems technology to assist dairy farmers would also contribute to the application of this technique to other domains in the agricultural area.

Chapter 4: Methodology

This chapter describes the AIM methodology which was developed in this research. Each process of the AIM methodology uses a set of techniques and representations in order to separate each process's inputs and results. The methodology consists of three main processes: analysis, interpretation and modelling (figure 4.1). Although elicitation is considered, in this thesis, as a process of the knowledge acquisition phase, the process was not part of the AIM methodology. The elicitation process was conducted previously by the knowledge elicitors of the dairy project (chapter 3). In the first process Analyze Knowledge, the knowledge elicited from the experts is analyzed using appropriate techniques. The results are input to the Interpretation Knowledge process. This process applies the KADS interpretation models to the analyzed knowledge in order to choose a template for the domain. In the last process, Model Knowledge, the final template is instantiated, resulting in the model of expertise of the domain area. The three processes of the methodology are described in more detail in order to explain how the techniques and representations were used.



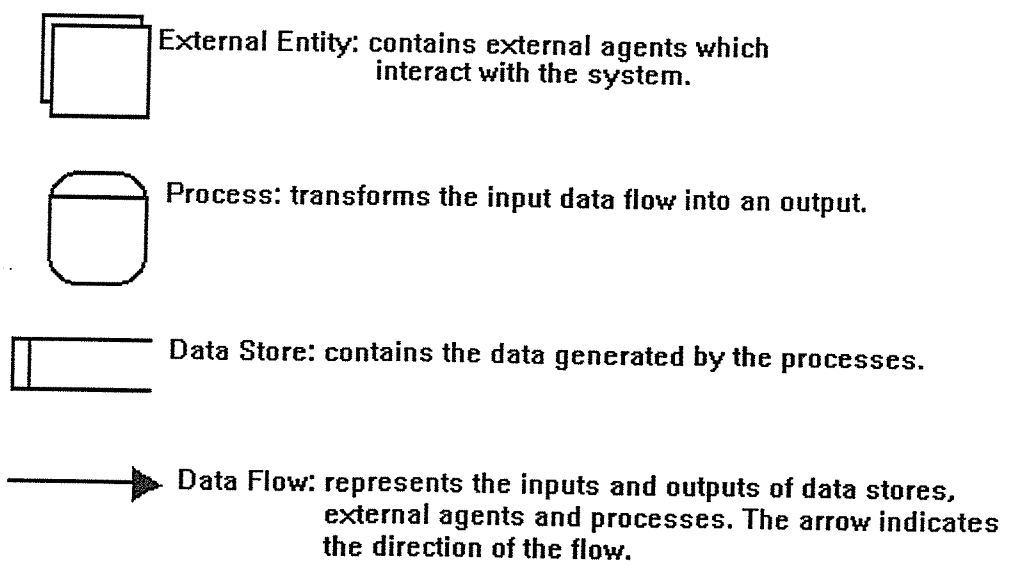


Figure 4.1. Data flow diagram of the AIM methodology.

4.1 The analysis process

The Analyze Knowledge process is expanded into two main processes: Analyze Transcripts and Create Graphic Representation and the Domain Knowledge data store into Domain Primitives and Network of Primitives (figure 4.2).

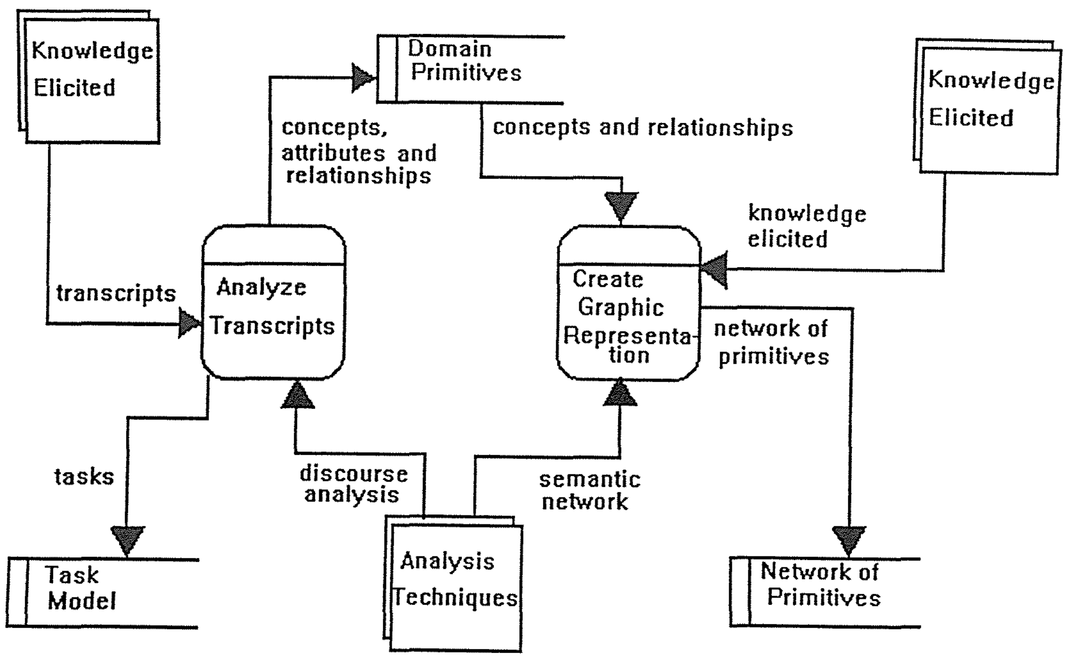


Figure 4.2. Techniques and representations of the Analyze Knowledge process.

In the Analyze Transcripts process, the data set obtained from the elicitation process is analyzed. The aim of this process is to refine the elicited data by eliminating redundant and unnecessary parts of knowledge and emphasising the important ones. To determine the meaningful parts of the elicited knowledge the discourse analysis technique is used. In this analysis the main concepts, relationships, attributes and functions of the domain are identified. The concepts, attributes and relationships form the Domain Primitives. The functions are used to construct the Task Model. These functions represent the overall tasks performed by the experts in order to solve/avoid problems and plan future actions.

The concepts and relationships obtained from the analysis are used as the input data for the second process which transforms these primitives into a graphical representation. To do that semantic networks are used. There are two uses for the semantic network technique. The first one is to graphically represent the relationships between concepts which can describe the connection between different concepts in a simpler way. The second use of semantic networks is to solve misunderstandings and/or disagreements between the experts. This can be done by constructing one network for each expert's main domain elements. These networks are then compared and integrated. The result is called the Network of Primitives. These are more general representations of the domain than the Domain Primitives, as the Network of Primitives does not necessarily contain all relationships between the domain concepts. Figure 4.3 shows an example of a Network of Primitives.

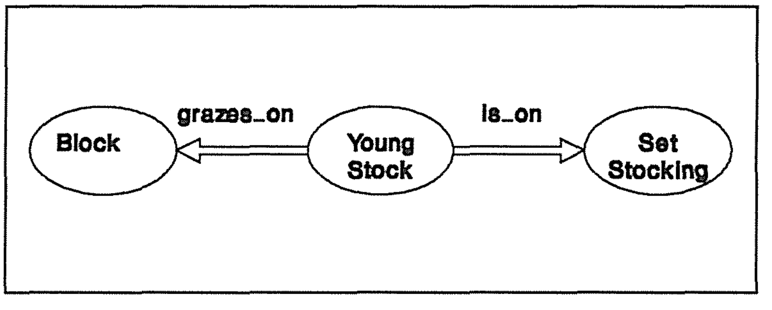


Figure 4.3. Example of Network of Primitives

4.2 The interpretation process

The expansion of the Interpret Knowledge process also resulted in two processes. The Domain Knowledge data store is again expanded into Domain Primitives and Network of Primitives (figure 4.4).

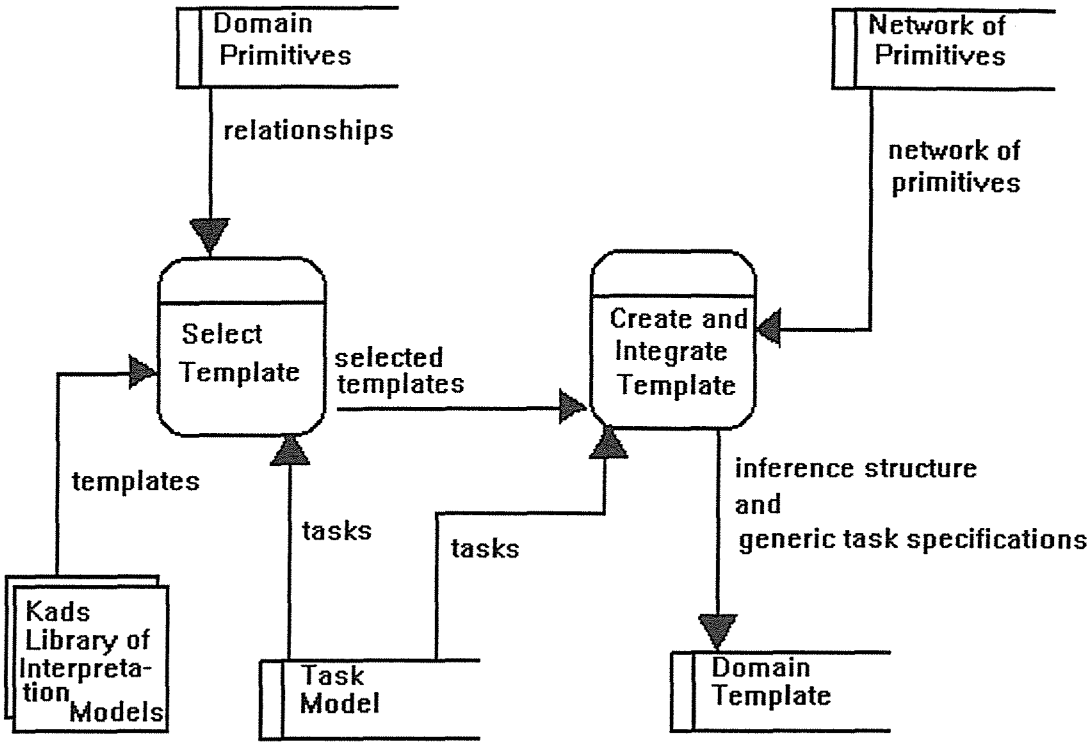
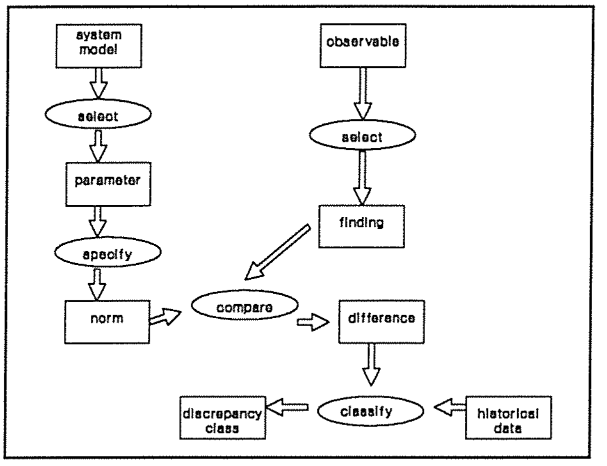


Figure 4.4. Techniques and representations of the Interpret Knowledge process.

In the first process Select Template, the interpretation models (templates) of the KADS library are analyzed in order to choose the template which best suits the domain features. The aim of this process is to choose a template to guide the modelling of the data. To select the template, the relationships of the Domain Primitives and the experts' functions represented in the Task Model are considered. If the domain area involves more than one problem-solving activity, more than one template may be chosen from the KADS library. Each template in the KADS library contains an inference structure and a generic task specification for a particular problem-solving activity. Figures 4.5 and 4.6 show an example of an inference structure and a task specification, respectively.



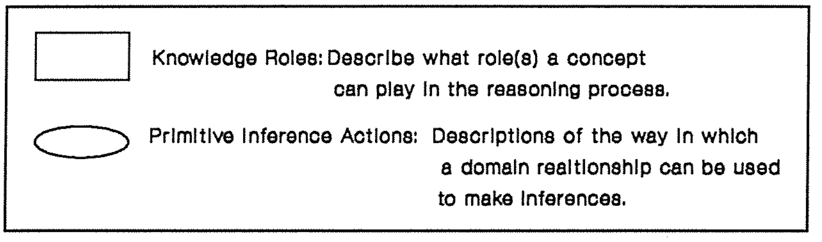


Figure 4.5. Inference structure of monitoring template (from Wielinga *et al.*, 1991).

In figure 4.6 below, the task structure represents the sequence of primitive inference actions in order to accomplish the monitoring task. The parameters of the inference actions are represented by the knowledge roles. The arrows indicate the transformation of an input knowledge role into an output.

task: monitoring

goal: execute a monitoring cycle in which the system activity acquires new data

control-terms: active-parameters = set of parameter

task structure

monitor(discrepancy)

select(system model, active-parameters)

do for each parameter and active-parameters

instantiate(parameter--->norm)

select(parameter--->observable)
obtain(observable--->finding)
compare(norm+finding--->difference)
classify(difference+historical data--->discrepancy)

Figure 4.6. Task specification of monitoring template (from Wielinga *et al.*, 1991).

The second process involves creating and integrating templates. It would be necessary to create a template when the problem-solving activity identified in the domain is not found in the KADS library. In a case where the domain knowledge presents more than one problem-solving activity, the integration of all activities may be necessary. This Create and Integrate Template process is mainly a customisation of template(s) of the KADS library according to the features of the domain. The result of this process is the Domain Template. This contains the inference structure of the domain and the generic task specifications. These specifications describe how the primitive inference actions and the knowledge roles of the inference structure can be combined to achieve a particular goal.

4.3 The modelling process

This process aims at developing implementation independent models of expertise. It is divided into three main activities according to the kind of knowledge to be modelled (figure 4.7).

The first process -Model Domain Knowledge- defines the knowledge in the domain layer. The domain layer consists of the set of Domain Primitives, Network of Primitives and a Data Dictionary. The Model Inference Knowledge is based on the Domain Template and the domain layer. During the modelling, the knowledge roles and primitive inference actions are instantiated with the domain elements. This process results in the inference layer which may involve revising the original inference template.

The last process Model Task Knowledge consists of the decomposition of the generic tasks specifications of the Domain Template into subtasks. The decomposition of all generic tasks constitutes the task layer.

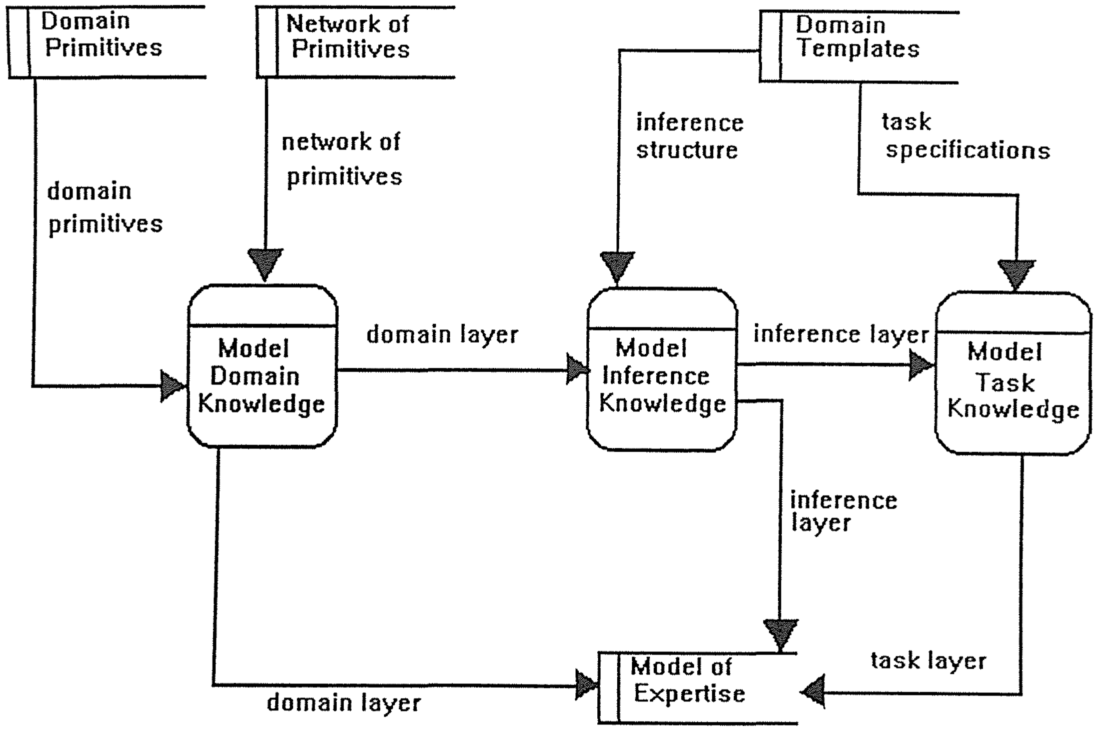


Figure 4.7.Techniques and representations of the Model Knowledge process.

4.4 Conclusion

The processes of this methodology are designed to be performed by the knowledge analyst. The role of the knowledge engineer played by the knowledge elicitor is not emphasised as the elicitation process is not described in detail. Analysis techniques are used to elaborate the raw data which is obtained from experts through elicitation techniques, such as interviews, by eliminating the noise material and identifying important concepts, relations and functions.

The data analysis led at the beginning of the AIM methodology to a bottom-up approach, with the data being studied in detail in order to build the domain elements. This approach was followed until the main domain elements were identified. Due to the large amount of data being considered in this research and the fact that the main domain elements had been identified as a first step, the bottom-up approach was then replaced by a top-down method. This started with the selection of the interpretation models (templates) which guided the interpretation and modelling of the data from a general template to a more detailed representation of the domain knowledge in the model of expertise.

Chapter 5: Case study and discussion

This chapter demonstrates the application of the AIM methodology. An example is used to illustrate how the data were analyzed, interpreted and modelled. The data set used in the application of the methodology was obtained from the dairy project. Results from two of the four experts interviewed by the members of the project team (knowledge elicitors), were chosen as the research data. Two sets of six transcripts representing the expertise obtained from the dairy farmers were analyzed, interpreted and modelled by the knowledge analyst. Since there is a large amount of data, only a section of one transcript will be used to illustrate specific points. The application of the methodology for all transcripts can be found in the Appendices. The main difficulties found in the application of the AIM methodology as well as the advantages and disadvantages of the techniques used will be discussed as the results are presented.

5.1 Analysis

The application of the AIM methodology started with the analysis of the data. It consisted of refining the data obtained by the participants of the Dairy Project during the semi-structured interviews with the experts.

5.1.1 Analyze Transcripts process

The analysis of the data started with the reading of the transcripts.
In total, 12 transcripts of an average of 15 pages each were analyzed. To illustrate how this was done, small sections of one transcript were chosen to be used as examples. The parts shown here were chosen as the important elements in the original transcript and were summarised in order to discuss the domain knowledge. The main concepts are underlined in the following summary of a transcript.

Example of a transcribed interview (9.1.92)

KE: Knowledge Engineer; EX: Expert (Dairy Farmer)

KE: What we want to get first is a description of how you see the current situation on your farm at the moment. Over the last few weeks what has happened? Having done that get an idea of what sort of planning horizon you are looking at the moment- how far you are looking. And then what level of planning you are doing- if it is formal or informal. The first thing is how would you describe the current situation of your farm?

EX: I feel I'm in a good position at the moment compared with other January, the rainfall we had through December and the last couple of weeks. I have not pasture scored at all but just what the cows are getting fed and the way they are milking, rotation length they are on, I could not be happier at the moment.

KE: Are you pasture scoring at the moment?

EX: Very informally, because dry matter is so low you can only really measure a straight height and the cows get to a point where they forage a lot harder. They are picking more than the plate meter can read.

KE: Are you monitoring your young stock?

EX: Yes, but very informally. At the moment they are ahead of target in terms of weight.

KE: Are sort of grazing system are they on?

EX: They are on the leased block next door.They are on set stock basically.

KE: What sort of rotation length are the dairy cows?

EX: They are on a 30 day round. I was trying to get it as long as possible and that is all the pasture cover and the feed requirements allowed me to do.

KE: In terms of Production, what levels of production have they been doing lately?

EX: At the moment 0.8 kg of milkfat/cow/day. This production is above the other years I'm quite pleased with it and i have made the decision to hold my rotation length to maintain the production level.

KE: Are you looking at cow condition at the moment?

EX: I'm looking at them every day. They score 4 1/2 at the moment. Good for this time of the year. They are being fed to increase their condition, but because they are milking they are just holding on.

KE: What kind of crop do you have at the moment?

EX: I have 5 acres of Winifred Rape which are almost grazeable. When the cows get to a point of 60% of their peak in litres or 0.6 of milkfat/cow/day or 12-13 litres/cow/day I will feed the crop to improve the production level.

KE: Have you identified any culls at the moment?

EX: I have just taken the cows off mating. I am still checking for empty cows. At the moment the options are 2-3 older cows which will definitely be sold.

5.1.1.1 The use of the discourse analysis technique

To analyze the transcripts, the discourse analysis technique was used. This analysis was based on the work of Kuiper & Kassirer (1987) and Belkin (1987). It was divided into two main steps. The first one consisted of reading the transcript and breaking it into short sections that corresponded to the main ideas expressed by the experts. Once these sections had been highlighted, the second step was to obtain from them the main concepts, attributes, relationships between concepts and actions (functions) performed by the experts. Two kinds of knowledge were obtained from the analysis. Declarative knowledge was obtained by focusing on correlated themes in the transcripts and ideas that had been emphasised by the experts. This kind of knowledge was represented by concepts, attributes and relationships. Procedural knowledge was obtained by analyzing the explanations given by the experts on how to perform a specific function (methods used, kind of knowledge applied etc).

5.1.1.2 Example of application of the discourse analysis technique

Following is an example of the domain elements which resulted from the application of the discourse analysis techniques to the sections shown above:

Concepts (In the order they were mentioned in the transcripts): The concepts represent the main topics mentioned by the experts. A concept can be an object, a method of evaluation used by the farmer, an element of the farm etc:

- **Pasture Scoring**: The act of measuring the amount of pasture on the farm. It is expressed in terms of quantity of dry matter/ hectare.

- **Dry Matter**: The amount of material left after water has been removed from pasture. **Attributes**: quantity, quality.

- **Plate Meter**: Tool used to measure the height of pasture cover.

- **Pasture Cover**: Vegetation produced by either a single or number of grass species, or combination of grass and clover species, generally regarded as being of long term or permanent contribution to the feed intake of grazing animals.

- **Grazing System:** Strategy under which pasture is harvested by the use of either animals or conservation practices.

- **Rotation Length:** This concept relates to a kind of Grazing System called Rotation System which is a planned succession of pasture use achieved by regularly moving the cows to pasture not recently grazed. The rotation length is the time taken for the cows to return to the same paddock having completed a "rotation" of the farm or block.

- **Block:** A separated farmland area managed as a unit with a specific purpose. The young stock is kept in this area.

- **Set Stocking:** Type of Grazing System where the animals are allowed to spread themselves over an area sufficiently large that they can support themselves there for lengthy periods without being moved to new pastures.

- **Cow Condition (also Body Condition and Condition Score):** the physical conformation of the animal in terms of muscle and fat cover in relation to its suitability for standing the winter and the calving process. **Attributes:** score (4 1/2).

- **Crop:** Supplement to be fed at strategic times of the year. It is usually planted between December and January to be fed in late February during the dry period when the pasture is not enough. **Attributes:** type (Winifred Rape).

- **Cull Cows:** Cows that are identified in the herd as possible options for removal from the herd. Many factors can influence the selection of cull cows: empty cows, bad mastitis cows, poor producing cows, old cows and thin cows.

- **Empty Cows:** Non-pregnant cow.

Relationships

In the following list of relationships, the numbers in parenthesis refer to the line numbers of the preceding transcript example. The relationships were obtained mainly from explanations about the farmers' work, i.e., what they did to solve problems, what kind of methods they used etc.

- The production level is an indicator of how the cows are being fed (8-9).
- At this time of the year dry matter is low (11).
- The young stock is kept on the block (18).
- The young stock is on set stocking (18).
- The dairy cows are on a 30 day rotation system (20).
- The production level is at the moment 0.8 kg of milkfat/cow/day (23).
- The rotation length can alter production level (24-25).
- Cow condition is 4 1/2 at the moment (27).
- It is difficult to increase cow condition when they are milking (28-29).
- The crop is Winifred Rape (31).
- The crop will be fed when the production level drop to 0.6 kg of milkfat/cow/day (32).
- The crop is fed to improve production level (33).
- Empty cows are being identified (35).
- Older cows are an option for culling (36).

Functions

The functions represent the activities the farmers perform in order to solve a problem, to evaluate the situation of the farm etc. They are tasks mentioned in the transcripts.

- Pasture score informally (11)
- Monitor young stock weight informally (15)
- Plan grazing system (20-21)
- Monitor cow condition informally (27)
- Monitor crop (31-32)
- Identify cull cows (35-36)

5.1.1.3 Results of the use of the discourse analysis technique

The domain elements obtained during the discourse analysis process generated the Domain Primitives and the Task Model (figure 4.2). The Domain Primitives consist of the set of concepts, attributes and relationships between concepts which were obtained during the analysis of all experts' transcripts. The Domain Primitives can be found in the Appendix 1. The Task Model resulted from the analysis of the functions obtained from the transcripts (figure 5.1). It represents the overall activities performed by the farmers in order to avoid/solve problems and plan future actions. A similar use of a Task Model can be found in Todd *et al.* 1993b.

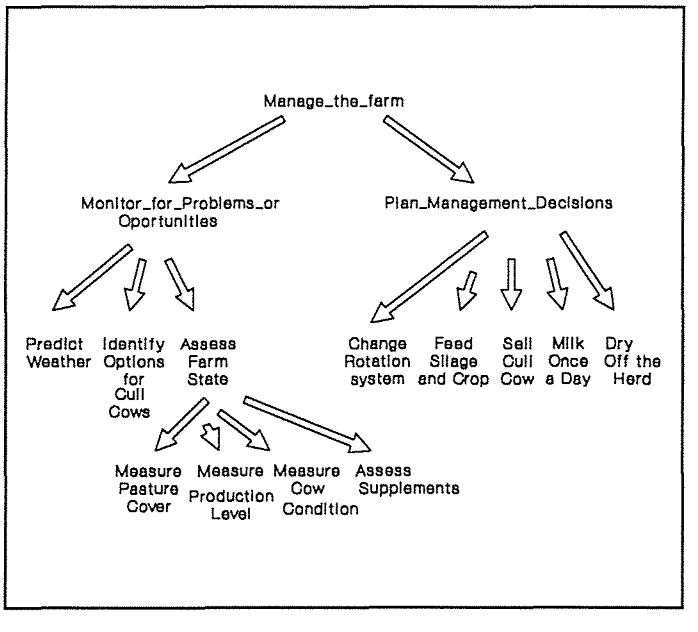


Figure 5.1. Task Model of the domain

5.1.1.4 Roles of a knowledge engineer

The participants of the dairy project who interviewed the farmers played the role of knowledge elicitors. They were responsible for obtaining the data and were directly in contact with the experts. The author of this thesis carried out the functions of the knowledge analyst: analysis, interpretation and modelling of data. The knowlegde analyst did not participate in the elicitation process. Instead, this person was responsible for the technical knowledge and for relating to the knowledge elicitors the unclear points in the transcribed interviews. The knowledge engineers were represented in the AIM methodology by different symbols as illustrated by figure 4.1. The knowledge elicitors were represented by the external entity Knowledge Elicited. The knowledge analyst was represented by the Analyze Knowledge, Interpret Knowledge and Model Knowledge processes. A knowledge engineer with a background in the domain area can be an advantage. In the application of elicitation techniques, an "expert" knowledge engineer has certainly more advantages than a "novice" knowledge engineer. Elicitation processes such as interviews can be more easily conducted by expert knowledge engineers since they know what kind of questions have to be asked without having to learn about the domain at the same time. Nonetheless, because of their high level of domain knowledge, such knowledge engineers may omit some questions during the elicitation process.

This was one of the problems in this analysis. Some of the basic concepts about the domain were not explained in the transcripts. This made the analysis of the interviews more difficult for a knowledge analyst without the same background as the knowledge elicitor. This was solved through research about the domain area as well as conversations with the knowledge elicitors about the transcripts. During these conversations the knowledge elicitors worked as auditors by verifying the domain elements obtained by the knowledge analyst during the application of the discourse analysis technique. Also, the bottom-up approach adopted by the knowledge analyst during the discourse analysis process, facilitated the comprehension and learning of the domain.

5.1.1.5 Practical considerations of the discourse analysis technique

The discourse analysis was a practical way of analyzing large amounts of data. Dividing the transcripts into smaller sections, made the analysis easier since every important concept was studied separately. Some of the difficulties in analyzing the transcripts were related more to the situation in which they were analyzed than the technique used. Analyzing the transcripts without having interviewed the experts may be more difficult since the transcripts do not express completely such psychological considerations as facial expressions or whether the expert was hesitant or vague about some points in the interview.

However there are advantages in having the elicitation and analysis processes conducted by different knowledge engineers. Missing and unclear knowledge was more easily identified by the knowledge analyst who had not been influenced by the elicited knowledge.

Another difficulty in analyzing the knowledge elicited was related to domain behaviour. The significance of the elements monitored by the farmers depended on the time of the year they were evaluated. This problem was solved by using the relationships to emphasise the event's importance as shown in the following example:

- <u>Pasture Cover</u> is not measured during summer because it is not high enough for the plate meter to read.
- <u>Pasture Scoring</u> is crucial during autumn/winter to make sure the pasture cover will be enough at calving time.

To consider this characteristic of the domain, the outputs of the discourse analysis were divided into two main management periods. The first period was called short term planning and comprised the management interval from January to March. These periods were characterised by different goals. The immediate goal of the first period was to maintain production at a satisfactory level. Because of this, elements like production level and cow condition were the priorities of the farmer, which was reflected in the domain concepts and relationships. This period was called short term planning because the farmers planned their actions in a two week horizon. The second period, from March until drying off date (usually May) was called long term planning, because the farmers planned activities aimed at a goal to be achieved in a longer term. The main goal of this period was to have a good pasture cover and cows in good condition in August during calving. There is not a specific date when the farmers change from short term to long term planning. The main indicator is a major increase in pasture growth rate, the autumn flush, which usually occurs between March and April. The autumn flush is the point when farmers change their goals from maintaining production level to building pasture cover. In order to take account of this time factor, the results of the analysis of the 12 transcripts were not integrated as one. To integrate the domain elements the semantic network technique was used.

5.1.2 Create Graphic Representation process

As mentioned earlier, the discourse analysis was a very practical technique for obtaining the domain elements, but it was still necessary to have a general picture of the domain which could show in a simpler way how the concepts were connected. To do so, the semantic network technique was used. The technique is a very natural way of representing the concepts and relationships in a graphical form. The domain elements were easily mapped from the Domain Primitives to the Network of Primitives considering concepts as nodes and relationships as the links between the nodes.

5.1.2.1 Examples of the use of semantic networks

Originally, the concepts and relationships of the domain were represented by one semantic network for each expert. These two networks were integrated resulting in a network of the domain. However, to make a clear graphical representation of the domain, this network was decomposed into several different networks. This set of networks was constructed in a hierarchical way. The first hierarchical level contained the main concepts and relationships. The other networks contained the decomposition of the main relationships and concepts. Figure 5.2 shows the high level network. The set of all networks can be found in the Appendix 2 (A2.2).

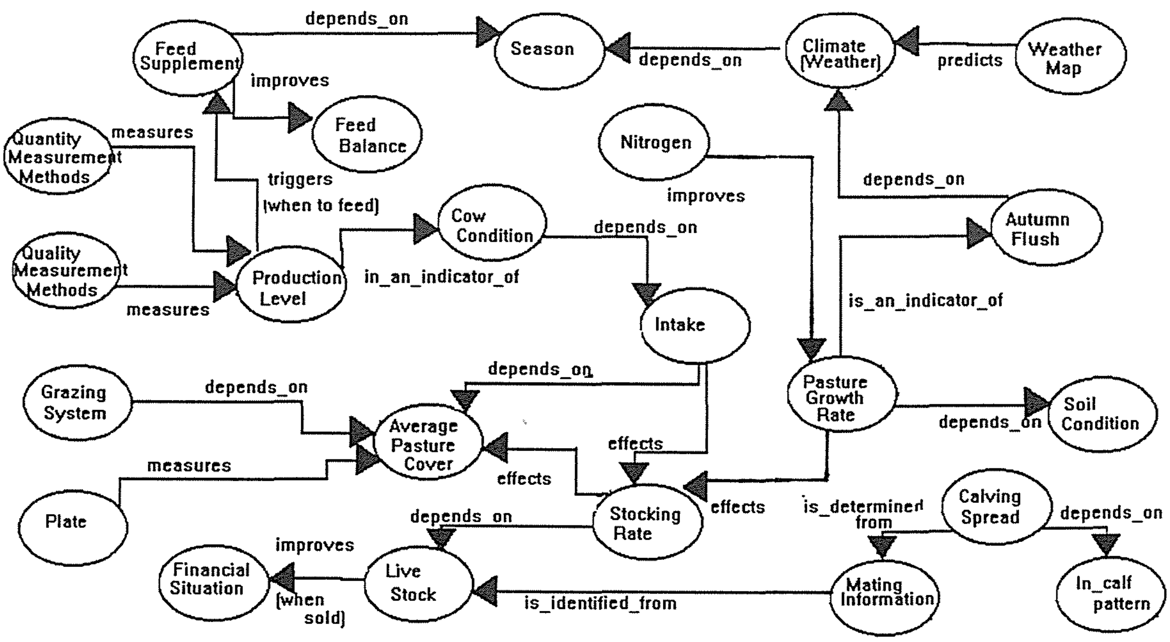
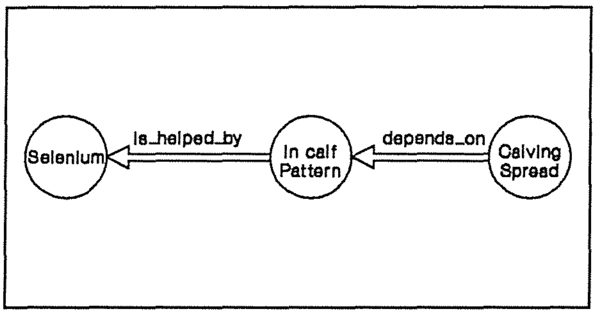


Figure 5.2. High level network of primitives

5.1.2.2 Practical considerations of semantic networks

As a knowledge representation, semantic networks were used to show the links between the concepts and relationships of the domain in a simpler way. As an analysis technique, they were used to compare and integrate the farmers' points-of-view. As mentioned before, the knowledge obtained from the two farmers was originally represented by two different networks. These networks did not present any major conflict. The main differences were the result of the importance attributed to various elements by the farmers. This was easily done by comparing the networks. Following is an example of how these differences were visualized and integrated (figure 5.3).

a)Expert 1:



b)Expert 2:

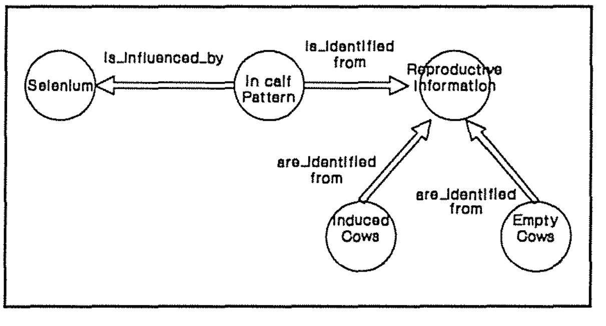


Figure 5.3. Different points-of-view represented by semantic networks

The comparison of the networks showed that expert 2 was more concerned and aware of the farm element "Reproductive Information" than expert 1. Expert 2 is more concerned about reproduction information because he has a higher stocking rate. Having a higher stocking rate requires the farmer to spend more time on prevention of empty cows and cows to be induced. This was identified from the comparison between the two networks. The analysis of both networks resulted in an integrated network where the knowledge of both experts is presented (figure 5.4).

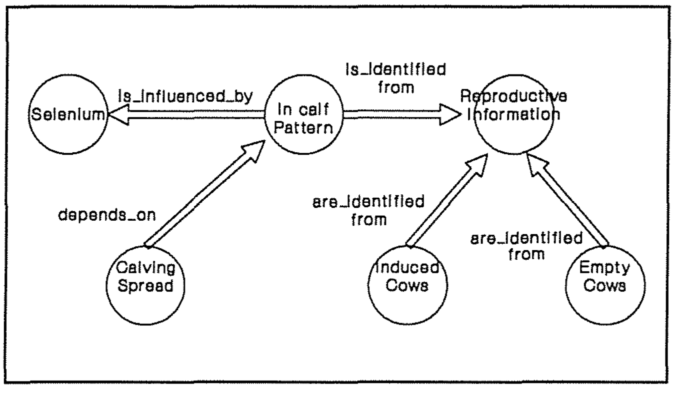


Figure 5.4. Example of integration of the experts' networks

Semantic networks were used here as a technique to clarify the meaning of the domain relationships (Jackson, 1990). Although the network may not be considered a very rigorous way of representing the domain knowledge, here the purpose was to use them as a general picture which allowed a faster understanding of the domain.

5.2 Interpretation

After analyzing the data, it was necessary to understand what kind of problem-solving behaviour was presented. This was done in the interpretation part of the AIM methodology which aimed at finding a template for the autumn-summer management domain. This problem-solving activity was chosen from the KADS library of interpretation models and used as a guide for the modelling process.

5.2.1 Select Template process

In this process, the Task Model and the relationships were used to choose the template(s) which would guide the modelling of the knowledge. To do so, the interpretation models of the KADS library were studied taking into account the characteristics of the domain. The selection of templates was not easy as the domain contained not one but several problem-solving activities. The problem was solved by choosing four interpretation models (figure 5.5). The first one was planning. This activity was the basis of the knowledge expressed by the experts since their main goal consisted of a plan of how to optimise milk production in the current season without putting the following season at risk. The second problem-solving activity was monitoring. This activity was frequently performed by the farmers since it was necessary to evaluate the farm's situation in order to make sure the plan was being followed. Diagnosis was another activity chosen due to the fact that the farmers had to identify causes of problems on the farm.

Finally, the last activity was prediction. The prediction of some events, such as the weather, helped the farmers to make more informed management decisions. To select the problem-solving activities from the KADS library, the main relationships obtained from the discourse analysis were considered. Following are some examples of the relationships which led to interpretation model selection:

.Monitoring:

* Monitor cow condition every day.

* Monitor production level every day.

.Prediction:

* Predict weather in order to make hay.

* Predict cow intake.

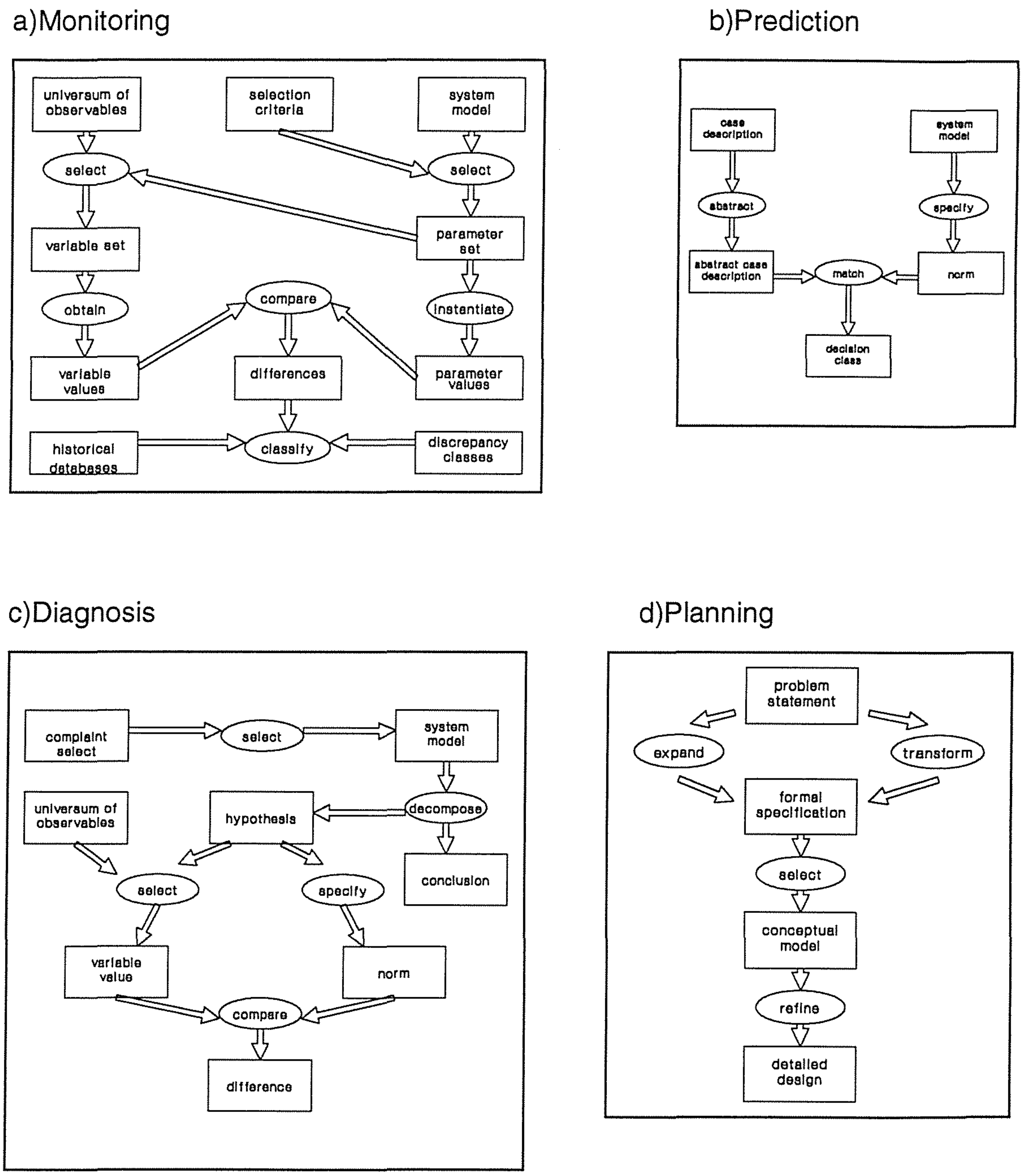


Figure 5.5. Inference structures (From Hickman *et al.*, 1989 and Breuker *et al.*, 1987).

.Diagnosis:

* Identify causes of low pregnancy rates.
* Identify cause of drop in production level.

.Planning:

* Plan rotation length to maximise pasture growth.
* Plan drying off date.

Because the KADS library did not propose an inference structure for the Planning activity, this was represented by the design inference structure. The design structure is considered here as the general template for the problem-solving activities which are classified as Synthesis activities (Hickman *et al.* 1989b).

5.2.2 Create and Integrate Templates process

In this process the domain template was created. The aim of the template was to guide the modelling of the data. The template contained two parts: the inference structure and the general task specifications.

5.2.2.1 Create and integrate inference structure

The four templates monitoring, diagnosis, prediction and planning were integrated into a single high level structure. This structure was constructed according to the functions of the Task Model. There were three main reasons for the integration of the inference structures mentioned above into a unique structure. The first was to eliminate some redundant knowledge roles and primitive inference actions such as "System Model" and "specify". The second was to have a structure which satisfied the domain characteristics in a more specific way removing some knowledge roles like "Abstract Case Description" which did not reflect the domain features. The last reason was to obtain an overview of what was happening in the domain.

5.2.2.2 Result of the integration of the templates

The result of the integration was an inference structure which, even though general, was closer to the domain features (figure 5.6). Because the inference structure of the domain is a general representation of the primitive inference actions, it covers both periods of the domain knowledge: the short term planning and the long term planning periods.

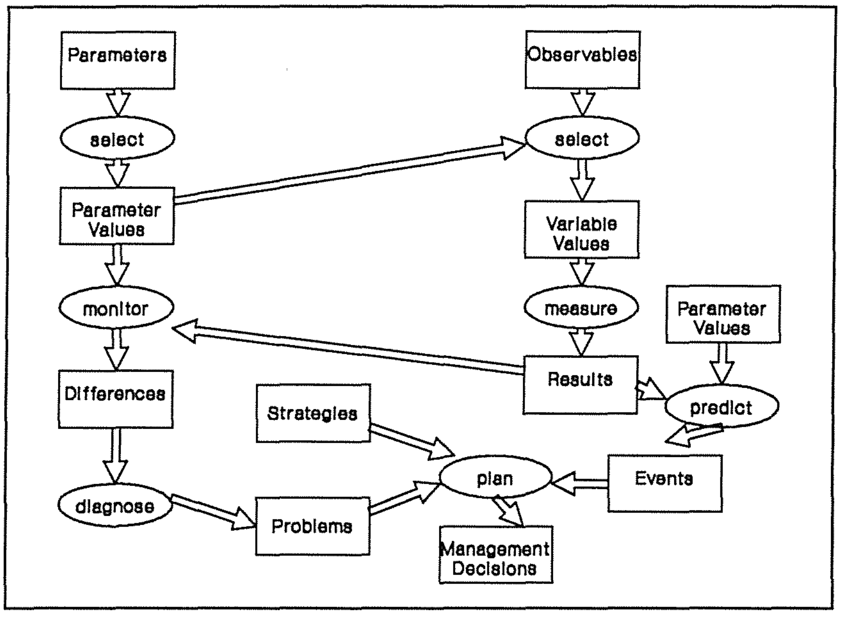


Figure 5.6. An initial inference structure for managing a dairy farm.

This inference structure represents the management of a dairy farm during the summer-autumn management period. Four activities are involved in this management process: monitoring, diagnosis, planning and prediction. To monitor the farm, the farmer would select an element(s), such as cow condition, pasture cover etc. These elements would be measured and the results compared to earlier results (results from other seasons) in order to check if the actual results matched with the expectations. This comparison might show a difference between the results. If this difference was significant, it would trigger the "diagnose" process to identify the cause of the difference.

To help the farmers make the best management decisions, some events on the farm would be predicted. Examples of such events are: weather, cow intake, average pasture cover etc. The decisions taken by the farmers would be planned according to the problems identified, the strategies designed to solve them and the prediction of the events.

Figure 5.7 shows the descriptions of the primitive inference actions and examples of knowledge roles of inference structure of figure 5.6:

select (Parameters): The inference structure starts with the selection of a number of parameters concerning the structure of the system being developed. The selection of parameters can be compared to an initial study of how an ideal system should work. This Primitive Inference Action also instantiates the parameters to their values.
input: Graph of optimum cow condition.
output: Expected cow condition.

select (Observables): Observables are selected from the real system. This selection may be guided by the parameter variables. The observables are transformed into variables through data abstraction.
input: Measurement Methods.
output: Method of condition Score.

measure: Before comparing the variables obtained from the system it is necessary to evaluate them. These Primitive Inference Actions transform the observables into results that can be compared to ideal parameter
input: Condition score of the cows.
output: Average condition score of 4.

monitor: The desired (parameter) values are compared with the result of the test and measure of the actual values. This comparison may generate no difference between the desired and actual values. But, in case a difference is produced, the problem causing the difference is identified.
input: Parameter Value of 5.0 and Result of 4.0 for the condition score.
output: Difference of 1.0 in condition score of the cows.

diagnose: This Primitive Inference Action classifies the difference to determine whether it is significant or not.
input: Difference of 1.0 in condition score of the cows.
output: Cow condition below target.

plan: Identifies the actions that have to be taken in order to overcome the problems and maintain the right results.
input: Problem: Cow condition below target.
Strategy: Feed supplement
output: Feed Supplement in order to improve cow condition.

predict: It uses some of the results to determine what will happen next in terms of events on the farm.
input: Results: -Pregrazing Pasture Cover of 1700 kg of dry matter/hectare
-Postgrazing Pasture Cover of 1000 kg of dry matter/hectare
output: Feed supplement in order to improve cow intake.

Figure 5.7. Knowledge roles and primitive inference actions

5.2.2.2 Create task specifications

Besides the inference structures, the templates in the KADS Library offered general descriptions of how the primitive inference actions and knowledge roles could be combined to achieve a specific goal. These descriptions are called task specifications (Breuker *et al.*, 1987). According to Wielinga *et al.* (1991), tasks represent fixed strategies for achieving specific goals. For this application, the overall task is the summer-autumn management of a dairy farm. The main goal associated with this task is to maximise current milk production without risking the next season's production. The following generic task shows the sequence of primitive inference actions in order to accomplish this goal (figure 5.8).

task **Manage_the_farm**

goal: Maximise current milkfat production without jeopardizing next season's yield.

task structure

select(Observable--->Variable Values)
select(Parameters--->Parameter Values)
measure(Variable Values--->Results)
monitor(Results and Parameter Values--->Differences)
diagnose(Differences--->Problems)
plan(Problems and Strategies----> Management Decisions)
predict(Results and Parameter Values--->Events)

Figure 5.8. Generic task specification of summer-autumn management of a dairy farm.

This generic task specification was used as a guide for the modelling of the task knowledge ratifying the top-down approach adopted. The generic task specification and the inference structure of the domain completed the Domain Template representation.

5.2.3 Practical considerations of the selection of templates

Various problems arose in practice when selecting templates. One of them was related to the complexity of the application which involved several parameters. Another problem was that there is no suitable template available for the planning activity. For this reason, the design template was used instead.

Finally, there was the problem of deciding when to select the templates. This was mainly due to the fact that the KADS methodology does not consider the selection of an interpretation model as a phase in the methodology, but as a tool to help in the knowledge acquisition phase (Breuker *et al.,* 1987). There is not a "right time" for the selection of the templates or how much data should be analyzed before being able to classify the problem-solving behaviour of the data in order to select the templates. The KADS methodology suggests that a reasonable amount of data should be analyzed before the selection which should be followed by the modelling of data. The AIM methodology is considering the selection and integration of templates as a whole process within knowledge acquisition rather than a tool. Although the process of selecting the template comes after the analysis of the data, since timing is not shown in data flow diagrams there is no problem with selecting a template before the discourse analysis has been completed. And because the templates are used as a guide only, they can be modified during the modelling process.

5.3 Modelling

The modelling of knowledge consisted of the construction of the model of expertise of the domain. Three layers of the KADS model of expertise were developed: domain, inference and task. The objective of the modelling of knowledge was to verify if the knowledge had been correctly represented as well as to combine these representations in order to construct the model of expertise.

5.3.1 Model Domain Knowledge process

The domain layer contained the declarative theory of knowledge: concepts, attributes and relationships. Some of the components of the domain layer resulted from the refinement of the results of the discourse analysis process. Others were transferred directly. The concepts and attributes obtained from discourse analysis were directly represented in the domain layer (Appendix 2 section, A2.1.1). A Data dictionary in which the main domain concepts were explained is also provided in the Appendix 2 (A2.1.2). The relationships were integrated and summarised (Appendix 2, section A2.1.4). The Network of Primitives created during the analysis was broken into several small parts according to the main points being addressed by each expert (Appendix 2, section A2.2).

5.3.2 Model Inference Knowledge process

In the "Model Inference Knowledge" process, the knowledge roles and primitive inference actions were instantiated using the domain layer in order to validate the inference structure. This resulted in the refinement of the inference template of the domain. This was decomposed into four different inference structures according to the four tasks identified in the domain: monitoring, diagnosis, planning and prediction (figure 5.8). Following is a description of the activities.

The monitoring activity starts with the selection of relevant parameters of the ideal system. In reality, this ideal system may be the historical data that the farmers keep from the earlier seasons. Using this data, the farmer can estimate an "ideal" behaviour for the animals as well as for the pasture system. The result of the application of this "select" primitive inference action is the parameter values which will be compared against the real values. These "real values" result from the selection of observables. Using the set of parameter values as a guide, the observables are selected from the real system which is, in this case, the dairy farm. These observables are the elements of the farm such as pasture growth rates, supplements, young stock, milking cows etc. The variable values are then measured. This measurement can be any kind of calculation regarding any element of the farm as, for example, condition scoring the milking cows. The result of the measurement is then compared to the parameter values. Figure 5.9a shows the part of the inference structure which represents the Monitoring activity.

The diagnosis activity is similar to the monitoring. The difference is that diagnosis has one more step which is to classify the differences which resulted from the comparison between the parameter and variable values (figure 5.9b). This classification consists of identifying the problem which is causing the difference between the ideal parameter and the results of the real system.

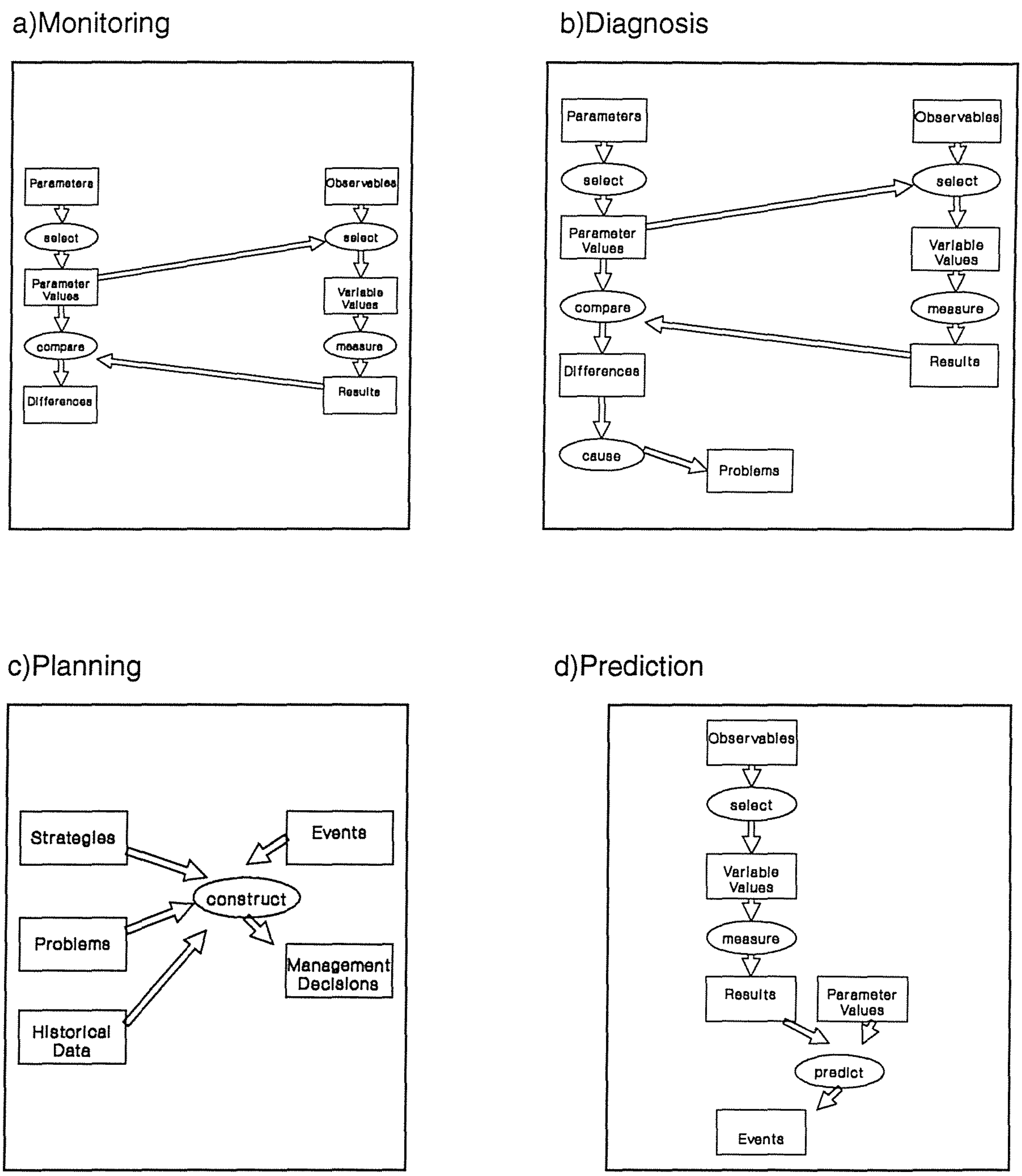


Figure 5.9. Parts of the inference structure of the summer-autumn management period of a dairy farm.

In the planning activity, the problems, strategies, historical data are used to construct a management decision (figure 5.9c). This construction is based on the combination of the strategies that the farmers have been using in earlier seasons in order to solve or prevent problems. To guide the planning of management decisions, a set of historical data containing information about the farm is used.

Finally, Prediction represents the last inference structure. In the same way as Monitoring, it also starts with the selection of Observables. The observables are then measured in order to obtain the Results. The difference between Prediction and Diagnosis is that the former does not require the Results to be classified as Problems. Instead, the Prediction activity uses the Results to predict the Events (figure 5.9d).

5.3.2.2 Example of an instantiation of the inference structure

To illustrate how the domain elements were instantiated with the inference structures, an example of the Prediction activity will be given. Given the relationship "Residual Pasture Cover predicts cow intake", the following activities would be taken by the farmer in order to predict cow intake:

Select Pregrazing and Postgrazing Pasture Cover as observables;
Measure quantity of grass dry matter of Pregrazing Pasture Cover;
Measure quantity of grass dry matter of Postgrazing Pasture Cover;
Use the results to predict cow intake.

5.3.2.3 Practical considerations of the Model Inference Knowledge process

Although they were seen as different processes in this methodology, the Create and Integrate Templates and the Model Inference Knowledge processes were strongly connected. In the former process, the Domain Template was constructed according to the domain features. In the "Model Inference Knowledge" process, the knowledge roles and primitive inference actions were instantiated with the domain layer in order to validate and refine the inference structure. This process resulted in the decomposition of the original inference structure into four components. Because of the difficulty in separating the selection of the template from the modelling of the inference knowledge, it was necessary to make clear the activities of each process. In the selection of the templates, the KADS library of interpretation models, the Task Model and the domain relationships were used as a guide to selection of the problem-solving activities. The activities were then customised according to the Task Model. This resulted in a general inference structure which ended the Interpretation process. The modelling of the inference knowledge started with the instantiation of the domain layer with the general inference structure (figure 5.6). This generated several refinements until a more detailed structure was achieved (figure 5.9). For this application, it was necessary also to decompose the inference structure as the domain knowledge presented more than one problem-solving activity.

However, in order to have an inference structure which could demonstrate all the activities of the domain in one structure, the four activities shown above were combined (figure 5.10). The Model Inference Knowledge process resulted in the inference layer, which consists of the inference structure and all concepts and relationships of the domain that can be instantiated with this structure (Appendix 3).

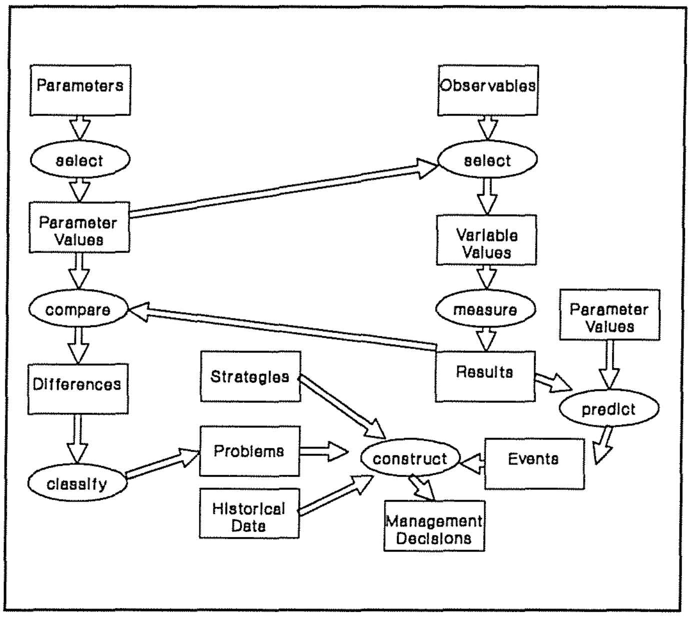


Figure 5.10. Final Inference structure of the summer-autumn management of a dairy farm

5.3.3 Model Task Knowledge process

This process was based on the generic task template and the inference layer. It aimed at refining the generic task template and decomposing the generic task into subtasks.

5.3.3.1 Instantiation and refinement of the generic task template

During the instantiation of the generic task template "Manage_dairy_farm" (figure 5.8) to the inference layer, the necessity of separating this generic task template into more activities was discovered. The template was divided into four activities according to the four inference structures created in the "Model Inference Knowledge" process. These four activities described the sequence of actions taken in order to monitor, diagnose, plan and predict the farm events. Figure 5.11 below illustrates the four activities which resulted from the generic task "Manage_Dairy_Farm" according to the inference structure of figure 5.10:

task **Monitoring Events**

goal: Make sure the results being obtained are according to the goals to be achieved.

task structure

select(Observable--->Variable Values)

select(Parameters--->Parameter Values)

measure(Variable Values--->Results)

compare(Results and Parameter Values--->Differences)

task **Diagnose Problems**

goal: Identify cause of problems

task structure

select(Observable--->Variable Values)

select(Parameters--->Parameter Values)

measure(Variable Values--->Results)

compare(Results and Parameter Values--->Differences)

classify(Differences--->Problems)

task **Planning Management Decisions**

goal: Organise a plan of actions that have to be taken in order to maintain production level.

task structure

construct(Problems,Strategies,Historical Data and Events----> Management Decisions)

task **Prediction of Events**

goal: Identify the future state of the events on the farm in order to take specific actions to avoid problems.

task structure

select(Observable--->Variable Values)

select(Parameters--->Parameter Values)

measure(Variable Values--->Results)

predict(Results and Parameter Values--->Events)

Figure 5.11. Decomposition of the generic task template.

The purpose of the task specifications was to show the sequence in which the primitive inference actions and the knowledge roles could be combined to achieve a goal.

5.3.3.2 Decomposition of the generic tasks

Although the generic task "Manage_Dairy_Farm" had been separated into four tasks, it was still necessary to describe these tasks in more detail. To do so, the four generic tasks of figure 5.11 were decomposed into subtasks according to the inference structure of figure 5.10. This generated a set of tasks, subtasks and goals associated with them. For each one of the four generic tasks: Planning Management Decision, Monitoring events, Diagnose Problems and Prediction of Events, several subtasks were created.

The decomposition of the tasks was based on the Task Structure of the KADS model of expertise (Wielinga *et al.* 1991) and involved three types of subtasks:

- Primitive Problem Solving Tasks: Tasks such as, "select(Production Level--->Production/cow/day)", specified as part of the domain template, i.e., primitive inference actions.

- Composite Problem Solving Tasks: Subtasks such as, "Measure Production Level", identified during the decomposition of the generic tasks.

- Transfer tasks: Tasks such as, "Provide total of litres" performed by an external agent which could be the expert or the system.

An example of a task decomposition is shown below (figure 5.12). It consists of the decomposition of the generic tasks Monitor Events and Diagnose Problems into one of their subtasks. In the examples, the type of task is written between parenthesis in order to illustrate the different types of tasks. One of the subtasks of the "Monitor Events" generic task is "Measure Production Level". It measures the production of milkfat and litres. One of the subtasks of the "Diagnose Problem" generic task is "Identify drop in Production Level" and it identifies a problem with the production level (in case there is one).

task: Monitor Events

subtask: Measure Production Level

goals: - Confirm that the animals are being well fed.

- Make sure the production goals are being achieved.

task structure:

select(Production Level--->Production/cow/day) (Primitive Problem Solving Task)

select(Production System--->Ideal Production) (Primitive Problem Solving Task)

Repeat every day

Provide total of litres (Transfer Task)

Provide kg of milkfat (Transfer Task)

Provide number of cows (Transfer Task)

Divide total of litres by number of cows (Transfer Task)

Divide total of milkfat by number of cows (Transfer Task)

measure(Milkfat/cow/day--->Results of Production Level) (Primitive Problem Solving Task)

compare(Ideal Production with Results of Production Level ---> Differences in Production Level) (Primitive Problem Solving Task)

End

task: Diagnose Problems

subtask: Identify drop in production level

goal: - Maintain Production Level

task structure

Measure Production Level (Composite Problem Solving Task)

If production has fallen below 0.6 kg of milkfat/cow/day

Then

classify(Difference in Production Level--->Drop in Production Level) (Primitive Problem Solving Task)

End If.

Figure 5.12. Example of decomposition of the generic tasks: Monitor Events and Diagnose Problems

5.3.3.3 Practical considerations of the Model Task Knowledge process

The classification of the subtasks into primitive problem solving tasks, composite problem solving tasks and transfer tasks was a major aspect in the modelling of the task layer. The primitive problem solving tasks were used as a guide for the development of the subtask structures. Each subtask should contain all the inference primitive actions presented in the generic task which originated it. For example, the generic task "Monitor Events" contained the primitive problem solving tasks: select, measure and compare. As a consequence all subtasks of "Monitor Events" would contain these primitive problem solving tasks in their structure. The composite problem solving tasks showed the connection between the subtasks. This can be seen in the example of figure 5.12. The task structure of the subtask "Identify drop in production level" contains the subtask "Measure Production Level", connecting the two subtasks. Finally the transfer tasks were the link between the subtasks and the expert. This is shown in the first example of figure 5.12 with the transfer tasks as, for example, "Provide total of litres".
The result of the decomposition of the generic tasks into subtasks and their interdependencies, was the Task Layer which can be found in the appendix 4 (section A4.1).

5.4 Conclusion

The result of the application of the AIM methodology was the model of expertise: domain, inference and task layers. The main characteristic of the domain layer of the model of expertise was the fact that it contained the static knowledge: concepts, attributes and relationships. The intention was to separate the problem-solving activities from the domain theory. This was achieved by separating the result of the discourse analysis process into two parts: the domain elements and the functions. These functions were the basis for the construction of the Task Model. This model contained the overall tasks performed by the farmer. Here, the model was constructed when a satisfactory amount of data had already been analyzed and the main functions identified.

The construction of the Domain Template was not an easy task. One of the difficulties was related to the integration of the problem-solving activities. To create the inference structure it was necessary to build several "versions" of the template. The modelling of the inference knowledge was a major aspect in the refinement of this inference structure. Instantiating the domain elements with the inference structure was a good way to check the completion of the inference structure. Also part of the template, the generic tasks were very useful to the construction of the task layer.

The task decomposition which originated the subtasks of the generic tasks was mainly based on the generic task template and the inference layer. In KADS, this decomposition is often very general and the subtasks do not differ much from the generic tasks of the library. Here, the tasks were decomposed to a more detailed level.

The result obtained with the application of the AIM methodology fulfilled the goals of this work. The overall objective was to model a complex domain which contained large amount of data. Besides this characteristic of the summer-autumn management domain, there was the fact that the knowledge analyst would have to learn about the domain without having participated in the elicitation process. The discourse analysis and semantic network techniques were a good way of dealing with the different levels of domain knowledge of the knowledge elicitor and the knowledge analyst. During the discourse analysis several points about the domain were clarified while the main concepts and relationships were obtained. As the final result of the application of the methodology, the model of expertise is the connection between the "raw" data obtained during the elicitation process and the implementation of the knowledge base.

Chapter 6: Conclusion

6.1 Results of the application of the AIM methodology.

The differentiation between elicitation and analysis processes was one of the main goals in the AIM methodology. By delineating these two activities, it was possible to separate the different kinds of knowledge which resulted from the two processes. The data obtained from elicitation was refined in the analysis process, resulting in a set of domain elements which did not contain redundant or non-important knowledge. These elements were the result of the application of the discourse analysis technique. Although a reasonable amount of time had to be spent in the analysis, the technique was a good way of analyzing the data while learning about the domain. The result of the application of the discourse analysis technique: concepts, attributes, relationships and functions were easily transferred to the domain layer of the model of expertise.

The domain elements were also easily mapped into the semantic nets. These representations were a good way of graphically representing the domain. However, semantic nets were not used here as a rigorous knowledge representation. They were used to illustrate the links between the relationships.

The function of the knowledge analyst was very different from that of the knowledge elicitors. It was possible to have a person with software engineering knowledge, the knowledge analyst, developing and applying the AIM methodology because the elicitation process was responsibility of the knowledge elicitors who had the domain knowledge. This was beneficial for both parties. The knowledge elicitors had the knowledge elicited (transcripts) checked by the knowledge analyst. On the other hand, the knowledge analyst had the knowledge elicitors' assistance while learning about the domain.

It was verified that, there is not a best time to choose the KADS interpretation model. In this methodology, the selection was made after the main concepts and relationships of the domain were identified. The conclusion was reached that an interpretation model of the KADS library should be selected once the main concepts are known.

The selection of the template(s) should be fairly easy if the domain was properly analyzed. After selecting the template(s), the rest of the analysis should confirm the selection.

It is clear that the KADS library is not complete, and some of the templates may need customisation depending on the complexity of the domain. Integration of the templates should be done after all the domain was analyzed. The templates may be revised during the modelling process, when the template(s) are being instantiated with the domain elements.

As a final result the model of expertise was a good representation of the different types of knowledge static and dynamic. The domain layer contained the static part of the knowledge while the inference and task layers contained the dynamic part.

6.2 Future work

Only a proportion of the transcripts obtained during the elicitation process was analyzed. It would be necessary to analyze the remainder of the transcripts to confirm the analysis of the summer-autumn management period. After this analysis is complete, the model of expertise could be used as the input for a prototype of the domain. The prototype would be used to confirm with the experts, knowledge elicitors, knowledge analysts and users that the concepts and functions of the domain were properly represented.

References

AAMODT, A. 1990. "A Computational Model of Knowledge-Intensive Learning and Problem Solving". In Wielinga, Boose, Gaines, Schreiber, van Someren (eds.). "Current Trends in Knowledge Acquisition". Netherlands: IOS Press, pp 1-20.

ALTHOFF, K. *et al.* 1990. "Multiple Knowledge Acquisition Strategies in MOLTKE". In Wielinga, Boose, Gaines, Schreiber, van Someren (eds.). "Current Trends in Knowledge Acquisition". Netherlands: IOS Press, pp 21-40.

ANDERSON, W. 1993. "Agriculture and Horticulture Production Systems I. 1993 Course Organisation and Administration.

BASDEN, A. 1989. "A Client-Centred Methodology for Building Expert Systems". Information Technology Institute, University of Salford. Salford, U.K.

BECKMAN, T. J. 1991. "Selecting Expert-System Applications". AI Expert, 6, n.2, pp 42-48.

BELKIN, N.J. *et al.* 1987. "Knowledge elicitation using discourse analysis". International Journal of Man-Machine Studies, 27, 127-144.

BOOSE, J. H. 1985. "A knowledge acquisition program for expert systems based on personal construct psychology". International Journal of Man-Machine Studies, 23, 495-525.

BRACHMAN, R. J. 1979. "On the Epistemological Status of Semantic Networks". In Findler N. V. (ed.). "Associative Networks". London: Academic Press; 3-50.

BRAMER, M. A. 1987. "Expert Systems in Britain: Progress and Prospects". In Bramer, M. A. (ed.). "Research and Development in Expert Systems III". British Information Society, pp 1-12.

BRATKO, I. 1989. "Fast Prototyping of Expert Systems Using Prolog". In Guida, G. and Tasso, C. (eds). "Topics in Expert Systems Design".Netherlands: Elsevier Science Publishers B.V.; 69-86.

BREUKER, J. *et al.* 1987. Model Driven Knowledge Acquisition: Interpretation Models. ESPRIT Project P1098 Deliverable D1 (task A1), University of Amsterdam and STL Ltd.

BREUKER, J. & Wielinga, B. 1989. "Models of Expertise in Knowledge Acquisition". In Guida, G. and Tasso, C. (eds). "Topics in Expert Systems Design".Netherlands: Elsevier Science Publishers B.V.; 265-295.

BRULE', J.F. & Blount, A. 1989. "Knowledge Acquisition". McGraw-Hill Publishing Company, 253 pp.

BRYANT, A. M. 1980. "Maximising Milk Production from Pasture". Proceedings of the New Zealand Grassland Association, 42,pp 82-91.

BULL, D. 1992. "Role of a Board director within New Zealand". New Zealand Dairy Exporter, April 92, 58-59.

BURTON, A.M. *et al.* 1988. "A formal evaluation of knowledge elicitation techniques for expert systems: domain 1". In Moralee, D. S. (ed.) Research and Development in Expert Systems IV, Proceedings of Expert Systems'87. Cambridge University Press; pp 136-145.

BYLANDER, T. & Chandrasekaran, B. 1987. "Generic tasks for knowledge-based reasoning: the "right" level of abstraction for knowledge acquisition". International Journal of Man-Machine Studies, 26, 231-243.

CAREY, J.M. & CURREY, J.D. 1989. "The Prototyping Conundrum". Datamation, June 1, pp 29-33.

CLANCEY, W.J. 1985. "Heuristic Classification". Artificial Intelligence, 27, 289-350.

CORDINGLEY, E.S. 1989. "Knowledge elicitation techniques for knowledge-based systems". In Diaper, D. (ed.). "Knowledge elicitation principles, techniques and applications". Ellis Horwood Limited, pp 89-171.

DAIRY FARMING ANNUAL. 1987. Summary of Statistics for the New Zealand Dairy Industry.

DAIRY FARMING ANNUAL. 1988. Summary of Statistics for the New Zealand Dairy Industry.

DAIRY FARMING ANNUAL. 1989. Summary of Statistics for the New Zealand Dairy Industry.

DAIRY FARMING ANNUAL. 1990. Summary of Statistics for the New Zealand Dairy Industry.

DAIRY FARMING ANNUAL. 1991. Summary of Statistics for the New Zealand Dairy Industry.

DAIRY FARMING ANNUAL. 1992. Summary of Statistics for the New Zealand Dairy Industry.

ERICSON, K.A. & Simon, H.A. 1984. "Protocol Analysis Verbal Reports as Data". The MIT Press, England, 462 pp.

FARM MONITORING REPORT, 1992. Farm Monitoring Report, North Central Region. Ministry of Agriculture and Fisheries. Policy Quit, Palmerston.

FREEMAN, P. 1985. "Knowledge Elicitation- A Commercial Perspective". Proceedings of the First International Expert Systems Conference (London 1-3 October 1985) 339-349.

GARG-JANARDAN C. & SALVENDY, G. 1987. "A conceptual framework for knowledge elicitation". International Journal of Man-Machine Studies, 26, 521-531.

GRAY, D.I. *et al.* 1992a. "A conceptual framework for an expert system to improve drying off decisions on seasonal supply dairy farms". Proceedings of the New Zealand Society of Animal Production, 52, pp 11-12.

GRAY, D.I. *et al.* 1992b. "The Potential Role of Expert Systems on Seasonal Supply Dairy Farms". DairyFarming Annual, 1992, pp 66-70.

GUIDA, G. & Tasso C. 1989. " Building Expert Systems: From Life Cycle to Development Methodology ". In Guida, G. and Tasso, C. (eds). "Topics in Expert Systems Design". Netherlands: Elsevier Science Publishers B.V.; pp 3-24.

HALL, T. 1991. "Big player on world trading stage". New Zealand Dairy Exporter, October 91, 40-41.

HART, A.1992. "Knowledge Acquisition for Expert Systems". Hart, A. (ed.), 196 pp.

HAYBALL, C. & BARLOW, D. 1990. "Skills Support in the ICL(Kidsgrove) Bonding Shop - A Case Study in the Application of the KADS Methodology". In Beery, D. and Hart, A. (eds). "Expert Systems - Human Issues", pp 85-97.

HAYES-ROTH, F., Waterman D. A. and Lenat, D.B. 1983." Building Expert Systems". Addison-Wesley Publishing Company, Inc, 444pp.

HAYWARD, S.A. *et al.* 1987. "Structured analysis of knowledge". International Journal of Man-Machine Studies, 26, 487-498.

HICKMAN, F. 1988. "What is KADS?". Report from Knowledge-based Systems Centre, Touche Ross Management Consultants. London.

HICKMAN, F. 1989a. "An Overview of The KADS Methodology". The Knowledge-Based Systems Centre of Touche Ross Management Consultants.

HICKMAN, F.R. *et al.* 1989b. "Analysis for Knowledge -Based Systems: A Practical Guide to the KADS Methodology". Ellis Horwood Limited, 190 pp.

HOLMES, C. W. and Wilson, G. F. 1984. "Milk Production from Pasture". Butterworths of New Zealand (Ltd), 319 pp.

HOLMES, C. W. 1993. "A Brief Description of New Zealand's Dairy Production System and Dairy Industry in New Zealand". Department of Animal Science. Massey University.

JACKSON, Peter 1990. "Introduction to Expert Systems" (2nd edition). Addison-Wesley Publishing Company; 526 pp.

JOHNSON, P.E. *et al.* 1987. "Specification of expertise". International Journal of Man-Machine Studies, 26, 161-181.

KAHN, G.S. & Bauer, M. 1989. " Prototyping: Tools and Motivations". In Guida, G. and Tass, C. (eds). "Topics in Expert Systems Design". Netherlands: Elsevier Science Publishers B.V.;47-68.

KARBACH, W. 1990. "Models of Problem-Solving: One label-One idea". In (Wielinga, B. et al eds). "Current Trends in Knowledge Acquisition". IOS Press; 173-189.

KELLY, G. A. 1955. "The psychology of personal constructs". New York:Norton, vol 1, 556 pp.

KEMP, E.A. & KEMP R.H. (1991). " The management of the system lifecycle in expert systems development". International Journal of Information Resource Management, vol 2, n.1.

KIDD, A.L. 1987. "Knowledge Acquisition". In Kidd A.L. (ed.). "Knowledge Acquisition for Expert Systems". New York : Plenum Press; 1-16.

KUIPERS, B. & Kassirer, J.P. 1987. "Knowledge Acquisition by Analysis of Verbatim Protocols". In Kidd, A.L. (ed.). "Knowledge Acquisition for Expert Systems". New York: Plenum Press; pp 45-70.

MATTOS, N.M. 1991. "An Approach to Knowledge Base Management". Springer Verlag; 247 pp.

McELDOWNEY, L. 1992. "Vertically integrated industry". New Zealand Dairy Exporter, May 92, 56-57.

MOTTA, E. *et al.* 1989a. "A Methodology and Tool for Knowledge Acquisition in Keats-2. In Guida, T. and Tasso C. (eds.). Topics in Expert System Design. Netherlands: Elsevier Science Publishers B.V.; 297-322.

MOTTA, E. *et al.* 1989b. "Methodological Foundations of Keats, The Knowledge Engineer's Assistant". In Wielinga, Boose, Gaines, Schreiber, van Somerer (eds.). "Current Trends in Knowledge Acquisition". Netherlands: IOS Press, pp 237-275.

MUSEN, M.A. *et al.* 1987. " Use of a domain-model to drive an interactive knowledge-editing tool". International Journal of Man-Machine Studies, 26, 105-121.

NEALE, I.M. 1990. "Modelling Expertise for KBS Development". Journal of Operational Research Society, 41, n.5, 447-458.

QUILLIAN, M. R. 1966. "Semantic Memory". In Semantic Information Processing. Minsky, M. (ed.). Cambridge, Massachusetts: Mit Press; 227-270.

ROBERTS, H.J. 1990. "Expert Systems Clubs: Design Methods". The Computer Journal, 33, n. 6, 556-561.

STEFIK, M. *et al.* 1983. "Basic Concepts for Building Expert Systems". In Hayes-Roth, F. ,Waterman D. A. and Lenat, D.B(eds). " Building Expert Systems". Addison-Wesley Publishing Company, Inc; 59-86.

TODD, E.G. *et al.* 1993a. "An Expert System to aid seasonal supply dairy farmers". The New Zealand Computer Society Conference. Auckland, New Zealand (in press).

TODD, E.G. *et al.* 1993b. "Using the KADS Methodology: Expert Systems in Farm Management". ANNES: The First New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems. Edited by Nikola K. Kasabov. pp 262-265.

WEEDS, B. 1993. "Dairy Board operational structure". New Zealand Dairy Exporter, June 93, 70-71.

WELBANK, M. 1990. "An Overview of Knowledge Acquisition Methods". Interacting with Computers, 2, 83-91.

WHITBY, M. 1988. "The Agricultural Handbook". Great Britain: Mackays of Chatham PCL, Chatham, Kent.

WIELINGA, B.J. *et al.* 1986. The KADS System Functional Description. ESPRIT Project P1098 Deliverable T1.1. Department of Social Science Informatics, University of Amsterdam, Herengracht 196, 1016 BS Amsterdam, The Netherlands.

WIELINGA, B.J. *et al.* (1987). "Knowledge Acquisition for Expert Systems". In Siekmann, J. (ed.). "Lecture Notes in Artificial Intelligence". pp 96-124.

WIELINGA, B.J. *et al.* 1991. "KADS : A Modelling Approach to Knowledge Engineering". ESPRIT Project P5248 KADS-II, University of Amsterdam.

Appendix 1: Results of the discourse analysis process.

A1.1 Domain Primitives

The domain elements obtained during the discourse analysis process generated the Domain Primitives and the Task Model. The Domain Primitives consist of the set of concepts, attributes and relationships between concepts which were obtained during the analysis of all experts' transcripts.

A1.1.1 List of concepts and attributes (in alphabetic order):

- Animal Health Problems
- Breed
- Calving Spread
- Cow Condition
- Cull Cows: empty cow, old cow, poor producing, mastitis cow.
- Dairy cow
- Dry Matter : quantity, quality
- Drying off
- Feed
- Feed Balance
- Feed Budget
- Feed Demand
- Feed Supply (Supplements): silage, hay, crop, urea)
- Grass Reproduction
- Grazing System: set stocking, rotational system
- Intake
- Leased Block
- Mating
- Milking Area
- Paddocks
- Pasture Cover: quality effect, seedhead

- Pasture Growth Rate
- Pasture Scoring: plate meter, measurement method
- Planning Horizons
- Postgrazing Pasture Cover (Residual)
- Pregrazing Pasture Cover
- Production Level: milkfat (kg), protein, litres
- Selenium
- Spore Count
- Reproductive Performance: in calf pattern, conception rate
- Rotation System : rotation length, paddock size
- Soil Temperature
- Somatic Cell Count
- Stocking Rate
- Zinc
- Weather (Climate Data): rainfall, hot dry wind, dryness, shower of rain
- Weather Map
- Wintering at home
- Young Stock: heifers, calves

A1.1.2 Relationships

The relationships obtained during the analysis of the transcripts of the two experts were not integrated in order to maintain the real-time characteristic of the domain. Instead they were divided into two main management periods.

A1.1.2.1 Short term period

The Short Term Planning Period (actions planned on a 2 week horizon) covers the period from January until March/April. This period is characterized mainly by the farmer's effort in maintaining production levels as high as possible. The farmers do not plan longer term actions because the events on the farm, during this period, can vary from one week to another.

Expert 1: Relationships from interviews of expert 1 while using short term planning:

- Interview 1 - 09.01.92 -

Current Situation of the farm:

- Rotation System: 30-day round.
- Production Level: 0.8 kg of milkfat/cow/day.
- Pasture Cover: 16-1700 kg of dry matter/hectare.
- Cow Condition: 4 1/2 +.

Relationships

Pasture Scoring

- Pasture is not scored formally until April.
- At this time of the year dry matter is very low.
- The plate meter only measures a straight height pasture cover.

Pasture Cover

- The situation of the Pasture Cover at this time of the year is very unstable.
- Pasture quality is not a problem at this time of the year.
- A good pasture cover situation leads to cows being well fed.
- Quality pasture may be evaluated by measuring milkfat and protein produced.
- Supplements are fed in order to allow pasture cover to build up.
- The Pasture Cover monitoring is an eye-assessment.

Climate Data

- Rainfall in December leads to a good pasture in the summer period.
- Climate data do not influence management decisions.

Rotation Length

- The rotation length is planned to maintain production level.
- The rotation length influences how well the cows are being fed.
- The milking cows are being kept in the paddocks during the night and during the day the cow are being fed the crop.
- The short term plan consists of rotation length, cull cows and supplements and covers the period of January to March.

Cow Intake

- Feeding the cows well is more beneficial than trying to have feed stocked up.
- The cows are drenched with zinc in order to avoid facial eczema.

Supplement

- Brassica and Rape are good options of crop.
- Silage is restrictive for the spring.
- The amount of supplements stocked depends on the summer.
- The crop growth depends on the summer.
- Hay and Silage are used mainly in Spring.
- The animals may be given silage to improve their feed balance.
- Silage is fed in early spring to get a balanced ration.
- Hay is fed in the winter.
- The crop is being fed during the day to complement the animals' feed.
- The supplements are fed in case the production level fall below 0.6 kg of milkfat/cow/day.

- The crop lasts 2-3 weeks.
- Silage is fed when the crop is gone.
- Supplements are fed when cow condition drops.
- Supplements are used when the rain causes a dramatic drop in pasture cover dry matter.

Feedbudget

- Feed Budget is not an issue at this time of the year.

Cull Cows

- Empty and older cows are the first options of cull cows.
- Poor producing cows are also considered of cull cows.

Reproductive Information

- Selenium in used to get higher pregnancy rates.
- The Calving Spread is planned using the animals' records which contain mating information of every cow.
- Empty cows are known after 3 weeks the bull is pulled out.
- Calves may bring dry cows to milk again.

Pasture Growth Rates

- Nitrogen may be used in the autumn to improve pasture growth rate.
- Pasture does not grow at a flat rate, but in bursts.
- It is necessary to take advantage of high pasture growth rates.
- The autumn flush is a time of high pasture growth rates.

Young Stock

- The young stock is monitored informally.
- The young stock is kept on the leased block.
- The young stock is on a set stocking.
- It is easier to manage the young stock than the dairy cows.

Production Level

- Production levels is an indicator of the condition of the cows.
- The lowest production level before feeding supplements is 0.6 kg of milkfat/cow/day or 12 litres/cow/day which represents 60% of the highest level.
- The peak of production is 22 litres/cow/day.
- Production level may drive the decision of feeding supplements and drying off the cows.
- Production level is firstly measured in litres, once those litres drop, the production is measured in terms of cow condition.

Cow Condition

- Cow condition is one of the main factors which influences the drying off process.
- If the cow gets to a condition of 3.5 or below, it will be sold or dried off.

Interview 2 - 23.01.92

<u>Current Situation of the farm:</u>

- Cow Intake: 9 kg of grass/cow/day
- Production Level: 0.7 kg of milkfat/cow/day
- Rotation System: 30-day round
- Cow Condition: 4 1/2.
- Number of cows: 138.

Relationships

Cull Cow

A cow is culled if its production gets to a level of 2 litres/cow/day.

Pasture Cover

- Pasture Quality is still not an issue, but quantity is.
- Pasture Cover and Pasture Growth are still monitored informally.

Rotation System

- The rotation length depends on the paddock size.
- The rotation system of the young stock is different from the rotation system of the milking cows.

Supplements

- Silage was not used so far.
- Silage may be fed after the crop.

Feedbudget

- A rough winter feedbudget will be done when the crop is finished.

Reproductive Information

- The reproduction information is got from the pregnancy test.

Young Stock

- The young stock is still on the block.
- The young stock does not come to the milking area until drying off time.
- The young stock is second priority to the milking cows.

Pasture Growth Rates

- The weather may increase pasture growth rates in case after a rainy period, the weather is not too hot or dry.

Cow Condition

- The lowest condition score acceptable is 3 1/2.
- It is hard to bring condition back once the cows have fallen to a low score.
- At this time of the year it is more beneficial to feed the cows well instead of saving feed.

Interview 3 - 13.02.92-

Current Situation of the farm:

- Production Level: 0.71 kg of milkfat/cow/day
- Cow Intake: 8 kg of grass dry matter and 2-3 kg of crop
- Rotation System: 27-day round
- Cow Condition: 4 1/2
- Number of Cows: 138

Relationships

Supplements

- The crop is being fed as a complement to pasture.
- The rain left the crop mud delaying the grazing.
- Crop may lose quality if left for too long to be fed.

Pasture Cover

- Pasture Quality is still not a problem.

Rotation Length

- Rotation Length was reduced from a 30-day round to a 27-day round because the crop was too muddy to be fed.

Cull Cows

- Six cows have been identified as empties.
- The empty cows will not be necessarily culled.

Young Stock

- The heifers are still being monitored informally.
- The heifers are above target weight.

Climate Data

- The weather which follows the rain is a major factor for the pasture growth rates.

- Interview 4 - 13.03.92-

Current State of the Farm:

- Production Level: 0.65-0.70 kg of milkfat/cow/day
- Cow Intake: 10 kg of grass dry matter
- Cow Condition: 4 1/2
- Pasture Cover: 1450 kg of dry matter/hectare
- Rotation System: 29-day round
- Number of cows: 138

Relationships

Young Stock

- Crop may be a problem for two year olds, because sometimes they don't eat it.
- The two year olds (Heifers) are being fed totally with grass at the moment.
- The one year olds (Calves) were fed with silage for one week in order to improve their feed.
- In terms of priority, the calves come first, then the dairy cows followed by heifers.

Production Level

- The methods of production level evaluation were so far: kg of milkfat/cow/day, litres/day.
- The milkfat production is still being measured at daily terms.
- The litres being monitored are the main indicator for changing the grazing system.

Cow Condition

- Cow condition is still monitored informally at daily basis.
- The cows got 3 1/2 before being milked once a day.

Pasture Scoring

- A formal measurement of Pasture Cover was done due to checking the actual situation before starting to feed supplements.

Pasture Growth Rates

- Growth rates are not formally assessed at this time of the year.

Pasture Cover

- Pasture Quality is still not an issue.
- The Residual Pasture Cover is an indicator of the pasture cover situation.
- A Pasture Cover situation of 1700 kg of DM/ ha may trigger a major evaluation.

Rotation System

- The rotation system was lengthened in order to feed the cows more pasture.
- Rotation Length dropped from 27 days to 24 days due to the rain which caused too much mud to feed the crop.

Supplements

- At this time of the year there are two options: Silage and Crop. The crop was chosen because it can deteriorate if not fed.
- Silage may be fed to the calves to allow the pasture cover on the block to grow.
- The Crop started on 31 January and finished 1st March.
- Rape crops are better than maize as its quality is proven by the production levels.
- Feeding only with crop may cause production level to drop. But it pays back in terms of growth rates.
- The herd was fully fed with the crop to take advantage of the rain to improve pasture cover.
- Silage is used when the herd has to gain weight.

Cull Cows

- The number of cull cows may be 16.
- So far there is eight empty cows.
- Some cows (one or two) will be carried over to get 160 cows in

the next year.

- The main reasons for culling are: empty cows, age, cell count, empty two year olds.

Herd Test

- The herd test is being used only for somatic cell count.

Climate Data

- Climate information still does not influence management decisions.

Drying off

- The drying off process starts by milking thin cows just once a day.

Reproductive Information

- Breeding records are a way of predicting number of empty cows.

Expert 2: Relationships from interviews of expert 2 while using short term planning

Interview 1 - 08.01.92 -

Current Situation of the Farm

- Production Level: 0.75 kg of milkfat/cow/day
- Cow condition: 4 1/2 - 5.0
- Rotation System: 21-day round
- Pasture Cover: 180 kg of dry matter/hectare

Relationships

Cow Condition

- Cows losing weight leads to drop in production level.

- Cow condition is assessed by a consultant to check if the evaluations done were correct.
- Monitoring cow condition is more necessary in the summer period.

Pasture Cover

- Pasture cover is monitored once a month at this time of the year.
- At this time of the year pasture quality is not a problem.

Supplements

- Warm and wet weather leads to crop growing very fast.
- The supplements are mainly used in the autumn/winter.
- Maize is used in autumn.
- The date for using the supplements depends on how long the milking is extended.
- Zinc is put in the system to balance high spore counts.
- Turnips have the disadvantage of growing very rapidly and getting to a point when they have to be all eaten, otherwise they are lost.
- The crops will be grazed in three week's time.

Empty Cows

- At this time of the year the empties are still with the herd being mated.

Climate

- Cold wet periods may cause low conception rates.
- The rainfall data are the only climate data recorded.
- The weather patterns are evaluated through weather maps.
- There are not too many decisions which can be done relying on weather maps.
- It is not possible to pick up long term provisions from weather maps.
- The wind is the worst factor for drying the ground out.

Reproductive Information

- The first reproductive information to be used is potential cows to be induced.
- It is better to have the whole herd pregnancy tested at this time of the year rather than to do a second round of inductions in September.

Cull Cows

- Heifers on production are not options for culling.
- Cull out producing cows is not so common.
- At the moment culling is not necessary.
- Empty and mastitis cows are the first options of cull cows.
- It is not worth to cull more cows than can be replaced.
- Culling is a way of making cash but is not a major consideration in the management process.

Production Level

- The lowest production level acceptable is 0.4 or 0.35 kg of milkfat/cow/day.
- The level of production is a function of the amount of feed on hand.
- Next season's production is not sacrificed for getting more milkfat at this season.
- Production level is the trigger to feed the supplements.
- Cow production is a better indicator of the Pasture Cover situation than pasture assessment methods at this time of the year.
- The speed of the drop of the production indicates how severe is the problem with the feed.

Young Stock

- The heifers are the lowest producers.
- The calves are kept separated from the milking cows.

- The calves graze on the block.
- The calves are kept separated until they are heifers.
- Heifers are more variable than milking cows.
- The yearlings do not affect any decision for the season.
- Young stock are weighed every six weeks.
- Young stock does not influence the drying off decisions unless they have to be brought to the milking area.

Stocking Rate

- High stocking rates demand larger amounts of feed.

Rotation System

- The rotation system is extended if the grass is not growing more than the requirements.

Feedbudget

- A feedbudget is not considered at this time of the year.

Drying off date

- External opportunities such as buying or renting more land make milking possible a bit longer.

Interview 2 - 22.01.92 -

<u>Current Situation of the Farm</u>

- Production Level: 0.66 kg of milkfat/cow/day
- Cow condition: 4 1/2
- Rotation System: 22-day round

Relationships

Pasture Cover

- Pasture cover was not measured before the interview.
- Pasture Cover and Production can be measured more accurately by per cow condition.
- The plate meter is only a guide of the Pasture Cover situation.
- Monitoring Pasture Cover gives a prediction of the cover ahead.
- Spore counts are not a problem at the moment.

Supplements

- The crop will be used in a week's time.
- If the crop is not in a tall enough state the grass silage will be brought first.
- The most important part of the planning at the moment is when to feed the supplements.
- Grass Silage is available for the summer.

Cow Condition

- Cow condition is dropping.
- A consultant will evaluate cow condition in the first week of February.

Pasture Cover

- Pasture Quality is still not an issue.
- If the pasture cover drops supplements are used to compensate the feed.

Reproductive Information

- A pregnancy test will be done early February.

- The pregnancy test is done six weeks after the bull is pulled off.
- The calving spread will be known in March.

Pasture Growth Rates

- Growth rates dropped off a bit.

Climate

- The weather map is still used for short term planning.
- Weather conditions and milking production are the main factors which will trigger the supplements to be fed.

Interview 3 - 12.02.92 -

Current Situation of the Farm

- Production Level: 0.67 kg of milkfat/cow/day
- Cow condition: 4 1/2
- Rotation System: 23-day round
- Pasture Cover: 1974 kg of dry matter/hectare
- Number of Cows: 321

Relationships

Drying off

- Two cows were dried off.

Production Level

- When production level drops to 0.6 silage will be fed.

Supplements

- The millet was first fed on February 1st.
- The crop had to be started because it was growing very fast.
- The maize silage will be used in autumn/winter.

Pasture Cover

- Quality of pasture cover is not an issue.
- The quickest way to build the cover is to leave more residual pasture cover.
- Feeding the crop is a way of building pasture cover.
- The possibility of getting 50 acres more is being evaluated.

Young Stock

- The heifers are holding condition (5 1/2).
- The calves are also holding condition.
- The weight of the young stock is a function of the breed.

Climate

- Weather maps may help in some decisions such as when to cut the hay.

Cull Cows

- It may be better to get rid of some cows to be induced in order to make the calving spread shorter.

Interview 4 - 11.03.92 -

Current Situation of the Farm

- Production Level: 0.62 kg of milkfat/cow/day

- Cow condition: 4
- Rotation System: 38-day round
- Pasture Cover: 2000 kg of dry matter/hectare
- Number of Cows: 319
- Cow intake: 10 kg of grass dry matter and 2 kg of silage/cow/day

Relationships

Production Level

- The drop in production was due to a slight drop in intake.

Cow Intake

- Cow intake is evaluated when the rotation length or supplements are changed.
- Residual Pasture cover is another indicator of cow intake.

Supplements

- The amount of silage fed is decided based on the rotation length aimed.
- The policy of feeding crop and let the grass grow higher can be done in a wet summer.
- Hay has not been touched yet.
- The maize silage may be fed to maintain and even to raise cow condition.
- The maize silage has more effect on cow condition than on milk production.

Rotation Length

- The rotation length was extended because the crop allowed an increase in the grass growth.

Pasture Growth Rates

- At this time of the year (late summer/early autumn) the grass would be regrown either using the crop or silage.
- The rainfall has not improved the growth rates as it should.

Young Stock

- For the heifers, the lowest condition score is 3 1/2 before drying off.
- A light heifer in April is a problem.
- The calves are weighting about 130 kg and this is not the best they could be.
- The heifers are maintaining the 5 1/2 condition score.
- The fact that the young stock will be brought home (to the milking area), can affect the drying off decision.

Cull Cows

- The low producers will not be culled yet because there is enough feed and supplements on hand.
- The first options of cull cows will be the empty cows.

Climate

- Weather maps are not being used as the hay is done.

Drying off

- What usually triggers the drying off decision is time of the year.
- The coming of the winter says that the drying off process is also arriving.

A1.1.2.2 Long term period

The Long Term Planning Period (actions planned for a 2-3 month's horizon) covers the period from March/April until May when the cows are dried off. This period aims to maintain a good Pasture Cover situation and cows in a good condition for the calving time.

Expert 1: Relationships from interviews of expert 1 while using long term planning.

Interview n. 5 - April 1992

Current situation of the farm:

- Cow Condition: 4 1/2
- Cow Intake: 9-10 kg of grass dry matter
- Pasture Cover: 1445 kg of dry matter/hectare
- Rotation System: 29-day round
- Production Level: 0.65-0.67 kg of milkfat/cow/day

Relationships

- The number of cows that will be wintered is very important when doing the feedbudget.

Cow Condition

- A condition score of 5 may not be the best at calving date.
- Between drying off date and calving date the cows can gain 0.5 in condition depending on the winter. Drying off the cows with a condition score of 5 may be uneconomical because they could have a condition of 5 1/2 or even 6 at calving time.

Cow Intake

- An intake of 10 kg is about right (9 is acceptable, but 12 is a waste).

Pasture Cover

- The Pregrazing Pasture Cover and Postgrazing Pasture Cover are used to calculate cow intake.
- The Pasture Cover is being monitored once a week.
- Pasture Quality is still not a problem.
- The Pasture Cover is used mainly for the feedbudget.
- 1400 kg of dry matter/hectare is the lowest acceptable level during this time of the year.
- The right pasture cover on early spring is about 2000-2200 kg of dry matter/ha.
- A Pasture Cover of more than 2200 kg of dry matter/hectare during early spring would be a waste of feed.
- A Pasture Cover below 2000 kg of dry matter/hectare might be risky and could lead to lower late production and cycling problems for the cows.
- At this time of the year the planning consists of monitoring pasture cover not allowing it to drop too much.
- The pasture cover is being monitored once a week to make sure that there will be enough pasture at calving date.

Rotation System

- The factors that can change the rotation system are: pasture cover, intake and cow condition.
- In early spring the rotation length is planned to maximise pasture growth.

Supplements

- Silage and Hay have not been used yet.

Cull Cows

- There are nine empty cows according to the pregnancy test.
- Two of the empties will be maintained as carryovers.
- Seven of the empties will be sent to the works.

Climate Data

- The weather can influence the animals' intake.

Young Stock

- Heifers and Calves are still being monitored informally.
- The calving date for the young stock is not the same as the milking cows.
- The young stock will be dried off earlier than the milking cows.

Pasture Growth Rates

- Pasture growth rates are worked out on ungrazed pastures.
- Pasture growth rates have been used to pick up the autumn flush.
- From April until calving date, growth rates and pasture cover are more important than production.

Drying off

- All animals are dried off at the same time.
- The drying off decision is influenced by the stocking rate, in the sense that a higher stocking rate means drying off early.

FeedBudget

- There are two dates for planning the feedbudget: until calving date and until feed supply equals feed demand at springtime.
- It is difficult to plan the feedbudget because of the variability of the spring feed supply.
- The feedbudget is considering a condition score of 41/2 + at

calving.

- Waiting for the autumn flush to trigger the winter feed budget.

Interview n. 6 - 1.5.92

Current State of the Farm:

- Production Level: 0.65 kg of milkfat/cow/day
- Number of cows: 129
- Rotation Length: 29-day round

Relationships

Cull Cows

- 9 culls due to : 6 empties, 2 on age, 1 on cell count.
- The culls went out of the system when the growth rate started declining.

Young Stock

- The calves are still being fed ad lib.

Pasture Cover

- It was decided to dry off when the pasture cover was about 1400.

Pasture Growth Rate

- The pasture growth rate which was going down considerably.
- The use of nitrogen can provide a longer milking time.
- In this case nitrogen was not applied because it would not reverse the trend of the pasture cover going below 1400.
- Nitrogen can be applied on autumn to prolong milking.

Reproductive Information

- Calving spread does not influence the feedbudget.

Drying off

- The reason why the good condition cows are dried off with the poor cows is that the fact that the poor cows are dried off does not improve the pasture growth rates because they still need to be fed as they were milking in order to gain weight.

Cow Condition

- If the cows are dried off early they can be at a condition of 4 and there is still time to gain condition until calving date. But if the drying off decision is delayed, the cows have to have a 4 1/2 condition at least, because there is not much time to gain condition until calving date.
- A bad decision in terms of the balance between feed supply and cow condition is done when the cows are fat and there is nothing to feed them in the spring.

Expert 2: Relationships from interviews of expert 2 while using long term planning.

Interview 5 - 05.04.92 -

<u>**Current Situation of the Farm**</u>

- Production Level: 0.59 kg of milkfat/cow/day
- Cow condition: 4
- Rotation System: 42-day round
- Pasture Cover: 1825 kg of dry matter/hectare
- Number of Cows: 319

- Cow intake: 8-9 kg of grass dry matter and 4-5 kg of silage/cow/day

Relationships

Production Level

- The Production dropped to 0.55 kg of milkfat/cow/day due to the windy weather.
- Silage was fed to improve the production.
- Now the production level is about 0.59 kg of milkfat/cow/day.
- Production is still an indicator whether the cows are being well fed.

Climate

- The windy weather last for ten days and was the reason for the drop in production by affecting how much the cows needed to eat.

Supplements

- Silage was bought in order to improve production.

Pasture Cover

- Pasture quality is still not an issue.

Rotation Length

- Rotation System was lengthened from a 36 to a 42 day round because the silage is being fed.

Reproductive Information

- The herd was pregnancy tested and the result was 64 empty cows, what represents about 20% of the herd.
- This percentage is far in excess of the results of last years. Some analysis is being done to see why this has happened.
- Selenium deficiency could not be the reason as it was used all

year.

- Induced cows will not be a problem.(15% of the herd will be induced which is an average percentage).

Cull Cows

- Because of the high number of empty cows, the poor producing cows that are in calf will be maintained.
- The main option of cull cows will be the empties.
- The plan is to cull 40 empty cows from the total of 64. They will be kept for more 3 or 4 weeks until the next herd test is done.

Young Stock

- The skinny heifers will be dried off if they are losing condition and there is not enough feed on hand.
- The skinny heifers will be dried off instead off culling poor producing cows.
- The two year old heifers are holding the 5 1/2 condition.
- The calves are weighting 145 kg what is about right.
- Young Stock will finish its grazing at the end of May and then will be brought to the milking area.

Interview 6 - 30.04.92 -

<u>Current Situation of the Farm</u>

- Production Level: 0.59 kg of milkfat/cow/day
- Cow condition: 4 1/2
- Rotation System: 60-day round
- Pasture Cover: 1825 kg of dry matter/hectare
- Number of Cows: 230
- Cow intake: 8-9 kg of grass dry matter and 6 kg of silage/cow/day

Relationships

Drying off

- 51 cows were dried off on 10 April.
- The skinny heifers were the first option of drying.
- The heifers were at a condition of 3 1/2 when it was decided to dry them.
- The reason for drying off the skinny heifers was that they would need a couple of months of very good feed to put their condition back. It is easier to put condition when they are not milking and if this decision was delayed until the end of the month there would not be enough time to put condition before calving.

Cull Cows

- 36 empty, 2 lame and 2 high cell count cows were culled.

Production Level

- Production has dropped from 0.58 to 0.55 kg of milkfat/cow/day in the middle of the month and now is back to 0.59-0.60.
- The drop in the production level was due to bad weather (cold, frosty).

Young Stock

- The condition of the calves has increased 10 kg during the last month.

Pasture Growth Rates

- Pasture growth rate is the critical factor at this time of the year in terms of the drying off decision.

Supplements

- The maize silage is being fed to complement the grass.

A1.2 Task Model

The Task Model (figure A1.1) resulted from the analysis of the functions obtained from the transcripts. It represents the overall activities performed by the farmers in order to avoid/solve problems and plan future actions.

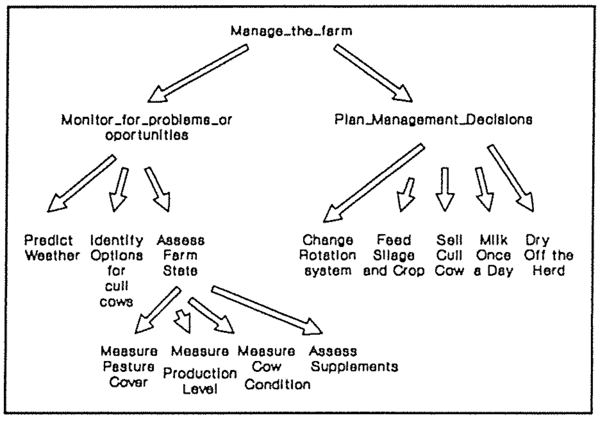


Figure A1.1. Task Model of the summer-autumn management of a dairy farm

A1.2.1 Functions obtained from the transcripts during the discourse analysis which led the construction of the Task Model

Functions obtained from the transcripts of Expert 1:

Short Term Planning:

- Pasture Score informally.
- Extend rotation length
- Measure production level
- Measure Cow condition
- Graze the crop
- Spore count every day
- Maintain the level of production at 0.8 kg/cow/day
- Pull the bull out
- Find out empty cows
- Complete the hay for the winter
- Use nitrogen to improve pasture growth rates

- Monitor (informally) the young stock
- Monitor rain.
- Predict cow intake.
- Do the herd test
- Drench cows with zinc
- Feed silage to the calves

Long Term Planning:

- Do the grass score weekly
- Do the pregnancy test for suspect cows
- Pull out induced and thin cows before calving
- Check the feed budget every week (for pasture growth rate)
- Milk the cows once a day
- Cut the pasture feed in half
- Bring silage to the system
- Calculate cow intake
- Regrass the crop
- Monitor the feedbudget according to the pasture growth rates.
- Milk the two year olds once a day (due to condition).
- Send culls to the works.
- Feed Silage .
- Dry off the cows.

Functions obtained from transcripts of Expert 2:

Short Term Planning:

- Measure pasture cover fortnightly.
- Measure cow condition.
- Weight young stock.
- Measure production level
- Measure grass growth fortnightly
- Monitor weather forecast every day

- Find out External Opportunities
- Put zinc in the herd's diet
- Monitor young stock feed.
- Analyze milking cow condition (by a consultant).
- Increase pasture cover residual.
- Monitor the silage to be fed
- Feed Crop
- Feed Maize Silage
- Sow Crop
- Extend Rotation Length to build pasture cover.
- Monitor Young Stock once a month informally
- Predict cull cows
- Drench calves
- Weight calves once a week

Long Term Planning:

- Measure Production Level daily
- Measure Cow Condition daily (informally)
- Measure Pasture Cover fortnightly
- Measure Pasture Growth Rates fortnightly
- Dry off skinny heifers
- Drench and Weight calves
- Feed silage
- Herd test
- Weight calves
- Monitor heifers
- Sell empty cows
- Dry off thin heifers
- Dry off milking cows

Appendix 2: Domain Layer

The Domain Layer consists of the Domain Primitives and the Networks of Primitives. This layer is the declarative part of the domain knowledge and so, does not contain problem-solving strategies. It consists of concepts, attributes and relationships.

A2.1 Domain Primitives

The Domain Primitives describe the set of the concepts, attributes and relationships between concepts. As a component of the domain layer, the Domain Primitives were transferred directly from the results of the discourse analysis. The concepts and attributes are the same and the relationships of the long term planning were integrated and summarised. A data dictionary is provided in order to describe the main concepts (Anderson, W. 1983; Holmes, C.W. and Wilson, G.F., 1987; Whitby, M. *et al.*, 1988).

A2.1.1 List of concepts and attributes (in alphabetic order):

- Animal Health Problems
- Breed
- Calving Spread
- Cow Condition
- Cull Cows: empty cow, old cow, poor producing, mastitis cow.
- Dairy cow
- Dry Matter : quantity, quality
- Drying off
- Feed
- Feed Balance
- Feed Budget
- Feed Demand
- Feed Supply (Supplements): silage, hay, crop, urea)

- Grass Reproduction
- Grazing System: set stocking, rotational system
- Intake
- Leased Block
- Mating
- Milking Area
- Paddocks
- Pasture Cover: quality effect, seedhead
- Pasture Growth Rate
- Pasture Scoring: plate meter, measurement method
- Planning Horizons
- Postgrazing Pasture Cover (Residual)
- Pregrazing Pasture Cover
- Production Level: milkfat (kg), protein, litres
- Selenium
- Spore Count
- Reproductive Performance: in calf pattern, conception rate
- Rotation System : rotation length, paddock size
- Soil Temperature
- Somatic Cell Count
- Stocking Rate
- Zinc
- Weather (Climate Data): rainfall, hot dry wind, dryness, shower of rain
- Weather Map
- Wintering at home
- Young Stock: heifers, calves

A2.1.2 Data dictionary

==

==

Name: Autumn flush

Description: Increase in pasture growth rate straight after increase in rainfall from March onwards.

Attributes: time, quantity

==

==

Name: Average Pasture Cover

Description: Average amount of pasture (kg dm/ha; kilograms of dry matter per hectare) on either a property or a defined block, paddock etc.

Attributes: quantity (kg DM/ha); quality (MJME/kg DM)

==

==

Name: Block

Description: A separated farmland area managed as an unit with a specific purpose. For dairy farms, the young stock is kept in this area.

Attributes: area

==

==

Name: Calf

Description: Young dairy cattle (up to about six months) not yet weaners.

Attributes: age; weight; breed.

==

==

Name: Calving Spread

Description: Period of time in which all animals should calve.

Attributes: number of days, date

==

==

==

==

Name: Carryover cow

Description: Cows which are not in calf and are kept milking for the whole season

Attributes:

==

==

Name: Climate

Description: The normal combination of radiation, temperature, moisture/precipitation, evaporation/transpiration, wind experienced by a location on an annual basis, with due regard to the usually occurring seasonal influences.

Attributes: rainfall, temperature, sunshine, wind

==

==

Name: Conception Rate

Description: Proportion of animals mated in the first four weeks of mating expressed as percentage. Usually it varies from 55 to 75%.

Attributes: percentage

==

==

Name: Condition Score

Description: Also Body Condition. The physical conformation of the animal in terms of muscle and fat cover in relation to its suitability for standing the winter and the calving process.

Attributes: Score

==

==

==

==

Name: Cover

Description: Describe the amount of herbage DM/ha in any one paddock.

Attributes: quantity; quality

==

==

Name: Crop

Description: Crop grown to provide supplementary feed for one or more classes of livestock at strategic times of the year. Crops may be fed in situ or mechanically harvested and fed to animals. Crops include Brassica species, grass species, maize and beets.

Attributes: species (See Feed Supplement for other attributes).

==

==

Name: Crossbreeding

Description: Mating genetically dissimilar males to females within a species of animals, such as different breeds, lines or strains.

Attributes: breed.

==

==

Name: Cull Cows

Description: Cows that are identified in the herd as possible options for removal. Many factors can influence the choice of a cull cow like (in order of priority): empty dairy cow,empty two year old, bad mastitis cow, poor producing cow, old cow and thin cow.

Attributes: See Livestock

==

==

==

==

Name: Dairy Farming

Description: A farming system where cattle is run with the purpose of producing milk.

Attributes:

==

==

Name: Dry Matter

Description: The amount of material left after water has been removed from a food. For example, herbage dry matter.

Attributes: percentage, quantity, quality.

==

==

Name: Drying Off

Description: Process of ceasing the lactation of dairy cows.

attributes: Date

==

==

Name: Empty Cow

Description: Non-pregnant cow.

Attributes: See Livestock

==

==

Name: Feed Balance

Description:A ration in which the nutrient ingredients are correctly proportioned for the needs of the livestock.

Attributes: nutrients

==

==

==

==

Name: FeedBudget

Description: The feed plan or budget is a simple balance sheet in which feed available is compared with feed required.

Attributes: feed supplement; pasture growth rate; livestock, average pasture cover; intake; cow condition; climate.

==

==

Name: Feed Supplement

Description: Other sources of feed supply besides the fresh pasture such as hay, silage and crops.

Attributes: Quantity (kg DM/ha); quality (MJME/kg of DM).

==

==

Name: Fertiliser

Description: Chemical or natural material applied to land to increase the supply of nutrient(s) from the soil.

Attributes: type; quantity.

==

==

Name: Grass

Description: The term is specifically used to describe the Poaceae family of plants which form the greater part of pasture species.

Attributes:

==

==

==

==

Name: Grazing System

Description: Strategy under which herbage grown on pastures are harvested by the use of either animals or conservation practices (or an association between both).

Attributes: type.

==

==

Name: Hay

Description: Hay is one of the options of Feed Supplement. It is made by mowing pasture, drying it while occasionally turning it in sunlight, then compressing it into bales.

Attributes: wastage; quantity (See Feed Supplement for other properties).

==

==

Name: Heifer

Description: Term used to describe a young female cattle. They describe a cow of up to and beyond her first calf.

Attributes: See Livestock

==

==

Name: Herd

Description: Name given to the set of bovine farm animals.

Attributes: See Livestock

==

==

==

==

Name: Herd Test

Description: A measure taken by a technician at regular intervals of individual dairy cows in a herd. The tests show the levels of milk constituents (milkfat, protein) for each cow as well as disease indicators (somatic cell).

Attributes: type

==

==

Name: Induced Cows

Description: Cows which would calve later than the majority of the herd. These cows are then injected with a hormone preparation which subsequently causes them to calve prematurely.

Attributes: See Livestock.

==

==

Name: Intake

Description: Amount of dry matter being consumed by an animal usually expressed in Kg DM/cow/day.

Attributes: quantity

==

==

Name: Livestock

Description: The number of farm animals. It includes the herd and the young stock.

Attributes: Liveweight, breed, condition score, intake, conception rate, number of animals.

==

==

==

==

Name: MA Cows

Description: Mixed Age Cows.

Attributes: see Livestock.

==

==

Name: Mastitis

Description: It is an inflammation of the mammary gland tissue usually due to a bacterial infection. The severity of the inflammation varies and two broad categories are recognised: clinical and subclinical mastitis. Mastitis affects milk composition which decreases the quality of the milk.

Attributes: type

==

==

Name: Mating

Description: The act of mating between males and females.

Attributes: number of animals; date.

==

==

Name: Mating rate

Description: Also submission rate. Number of cows being mated in the first four weeks of mating. In New Zealand conditions, it varies from 70 to 95% of the herd.

Attributes: percentage

==

==

Name: Mating Record

Description: Information about the date the animals were mated and which animals are expected to be in calf and which animals are expected to be empty.

Attributes: cow number, mating date.

==

==

==

Name: Nitrogen

Description: Chemical element usually used as fertilizer for increasing the rate of growth of pastures at different times of the year.

Attributes: quantity, quality

==

==

Name: Paddock

Description: Unit of division of farms. Correspond to an area of pasture that is used as fundamental unit in grazing management plans. It is dependent upon the size of the herd and grazing management standards.

Attributes: size

==

==

Name: Pasture

Description: Herbage produced by either a single or number of grass species, or combination of grass and clover species, generally regarded as being of long term or permanent contribution to the feed intake of grazing animals.

Attributes:

==

==

Name: Pasture growth rate

Description: Rate of growth of pasture measured as Kg DM/ha/day.

Attributes: quantity

==

==

Name: Pasture Scoring.

Description: The act of measuring the pasture cover on the farm. It is expressed in terms of quantity of dry matter/hectare.

Attributes: Method

==

==

==

Name: Plate Meter

Description: Tool used to measure the height of the pasture cover.

Attributes:

==

==

Name: Pregnancy rate

Description: Also in calf pattern. Number of cows which successfully become pregnant during the first four weeks of mating. It is a relationship between mating rate (MR) and conception rate (CR): PR = MR * CR

Attributes: percentage

==

==

==

Name: Pregnancy Test

Description: Testing by physical, electronic or chemical means for the presence of a foetus in a mated female.

Attributes: date, number of animals tested.

==

==

Name: Pregrazing Pasture Cover

Description: Amount of pasture (kg of DM/ha) on a defined block or paddock before being grazed.

Attributes: quantity (kg DM/ha); quality (MJME/ kg DM)

==

==

Name: Production Level

Description: Measurement of the amount and quality of milk produced.

Attributes: quantity, quality.

==

==

==

Name: Reproductive Information

Description: General information which results from pregnancy testing the cows like: number of empty cows, conception rate, date of mating and cows to be induced.

Attributes:

==

==

Name: Residual Pasture Cover.

Description: Also Postgrazing Pasture Cover. The amount of pasture or herbage dry matter remaining after grazing.

Attributes: quantity, quality.

==

==

Name: Rising

Description: Term used to describe the animal about to become a certain age e.g., a rising two year old is an animal approaching two years of age.

Attributes:

==

==

Name: Rotation Length

Description: The time taken for the stock to return to the same paddock having completed a "rotation" of the farm or block.

Attributes: number of days.

==

==

==

==

Name: Rotational System

Description: A planned succession of pasture use achieved by regularly moving mobs of animals to pasture not recently grazed. The pasture area is divided into several pieces: the paddocks.

Attributes: number of days, number of paddocks, number of animal in each paddock.

==

==

Name: Selenium

Description: Chemical element used as mineral supplement to the cows' diet which is very important for ensuring good pregnancy rates.

Attributes: quantity, quality

==

==

Name: Set Stocking

Description: The term applies to grazing livestock, where animals are allowed to spread themselves over an area sufficiently large that the animals are supported there for lengthy periods without being moved to new pastures.

Attributes: size of the area; number of animals.

==

==

Name: Silage

Description: Pasture having been allowed to grow from grazeable height (5-10 cm) to about 20 cm, is harvested green, being dried only minimally, and is compacted in a stack, with air excluded. Under anaerobic conditions the material is preserved for feeding to livestock.

Attributes: wastage (see Feed Supplement for other properties).

==

==

==

==

Name: Soil Fertility

Description: The supply of nutrients available (to plants) from the soil. May be natural or induced by application of fertiliser or manure.

Attributes:

==

==

Name: Somatic Cell Count

Description: Count of the number of white blood cells in the milk. This gives an indication of sub-clinical mastitis.

Attributes: percentage of cells.

==

==

Name: Species

Description: Subdivision of organisms- defined as the ability to produce viable offspring e.g. Orange and Apple.

Attributes:

==

==

Name: Spore Count

Description: It is the count of the fungal spores contained in the pasture eaten by the animals. These fungal spores contain a toxin which damages the animal's liver leading to jaundice, photosensitivity of the skin, and general weakness.

Attributes: quantity

==

==

Name: Stocking Rate

Description: The number of animals carried or run on a defined area of land.

Attributes: number.

==

==

==

==

Name: Weather

Description: The amount of radiation, temperature, precipitation, evaporation/transpiration and/or wind experienced in a specific time period. In the long term weather contributes to climate but in the short term may differ markedly from the norm.

Attributes: rainfall, temperature, wind, sun

==

==

Name: Weather Map

Description: Two dimensional representation of the state of the air pressure at a specific height above the surface of the earth at a specific time.

Attributes:

==

==

Name: Weed

Description: Plant growing in a place or at a time when it is of no use or has a detrimental effect.

Attributes: type.

==

==

Name: Young Stock

Description: Also replacement stock. Animals which are kept in the block separated from the milking cows. They demand a different grazing system and feed from the milking cows. The young stock usually consists of rising one year old animals (Yearlings) and rising two year old animals (Heifers)

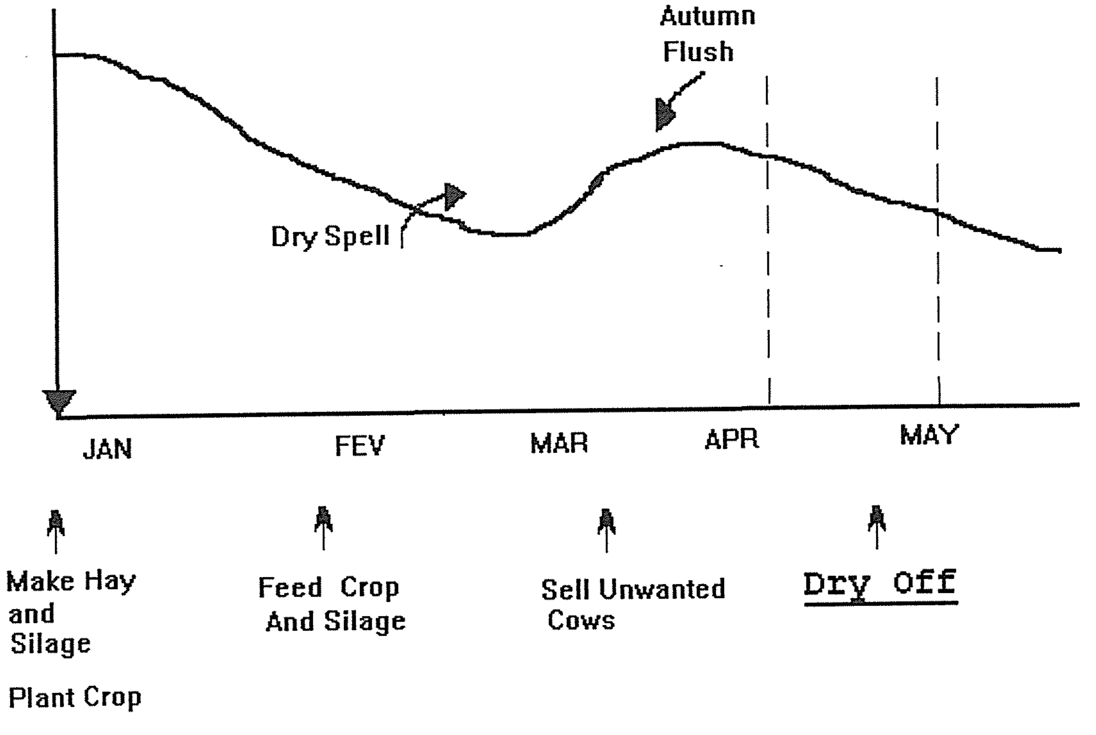
Attributes: See livestock.

==

==

A2.1.3 Graphical representation of the summer-autumn management period.

This diagram represents the events on a dairy farm during the summer-autumn period according to standard pasture growth rates. The main goal of this period is to maximise current milkfat production without jeopardizing next season's production.



The summer-autumn period is divided into two intervals according to the goals to be achieved:

Short term planning: Jan-Mar/Apr

The main goal of this period is to maintain production levels as high as possible.

Long term planning: Mar/Apr-May

The main goal of this period is to make sure there will be enough feed (pasture and supplements) during calving time.

A2.1.4 Relationships

The relationships obtained during the discourse analysis process were integrated and summarised. Following are the main relationships which characterise the short term and long term planning for the two farmers interviewed.

Short term planning

- At this time of the year dry matter is very low.
- A good pasture cover situation leads to cows being well fed.
- Quality pasture may be evaluated by measuring milkfat and protein produced.
- Supplements are fed in order to allow pasture cover to build up.
- Rainfall in December leads to a good pasture in the summer period.
- The rotation length is planned to maintain production level.
- The rotation length influences how well the cows are being fed.
- Feeding the cows well is more beneficial than trying to have feed stocked up.
- Crop growth depends on the summer.
- The animals may be given silage to improve their feed balance.
- The supplements are fed in case the production level fall below 0.6 kg of milkfat/cow/day.
- Silage is fed when the crop is gone.
- Supplements are fed when cow condition drops.
- Nitrogen may be used in the autumn to improve pasture growth rate.
- The autumn flush is a time of high pasture growth rates.
- The young stock is kept on the leased block.
- The young stock is on a set stocking.
- Production level is an indicator of the condition of the cows.

- Production level may drive the decision of feeding supplements and drying off the cows.
- Cow condition is one of the main factors which influences the drying off process.
- The reproduction information is got from the pregnancy test.
- Crop may lose quality if left for too long to be fed.
- The weather which follows the rain is a major factor for the pasture growth rates.
- Crop may be a problem for two year olds, because sometimes they don't eat it.
- The methods of production level evaluation are: kg of milkfat/cow/day, litres/day.
- The amount of litres measured are the main indicator for changing the grazing system.
- The cows got 3 1/2 before being milked once a day.
- Pasture quality is still not an issue.
- The residual pasture cover is an indicator of the pasture cover situation.
- A pasture cover situation of 1700 kg of DM/ha may trigger a major evaluation.
- The main reasons for culling are: empty cows, age, cell count, empty two year olds.
- Cows losing weight leads to drop in production level.
- High stocking rates demand larger amounts of feed.
- The rotation system is extended if the grass is not growing more than the requirements.
- Monitoring pasture cover gives a prediction of the cover ahead.
- The quickest way to build the cover is to leave more residual pasture cover.
- Residual pasture cover is another indicator of cow intake.

Long term period

- The number of cows that will be wintered is very important when doing the feedbudget.
- An intake of 10 kg is about right (9 is acceptable, but 12 is a waste).
- The Pregrazing pasture cover and postgrazing pasture cover are used to calculate cow intake.
- The pasture cover is being monitored once a week.
- Pasture quality is still not a problem.
- The pasture cover is used mainly for the feedbudget.
- A pasture cover below 2000 kg of dry matter/hectare might be risky and could lead to lower late production and cycling problems for the cows.
- The factors that can change the rotation system are: pasture cover, intake and cow condition.
- Pasture growth rates have been used to pick up the autumn flush.
- From April until calving date, growth rates and pasture cover are more important than production.
- The drying off decision is influenced by the stocking rate, in the sense that a higher stocking rate means drying off early.
- The use of nitrogen can provide a longer milking time.
- The reason why the good condition cows are dried off with the poor producing cows is that the fact that the poor cows are dried off does not improve the pasture growth rates because they still need to be fed as they were milking in order to gain weight.
- If the cows are dried off early they can be at a condition of 4 and there is still time to gain condition until calving date. But if the drying off decision is delayed, the cows have to have a 4 1/2 condition at least, because there is not much time to gain condition until calving date.

- A bad decision in terms of the balance between feed supply and cow condition is done when the cows are fat and there is nothing to feed them in the spring.
- Production is still an indicator whether the cows are being well fed.
- Pasture growth rate is the critical factor at this time of the year in terms of the drying off decision.

A2.2 Network of Primitives

The Networks of Primitives represent the connection between the concepts and the relationships. They do not contain all the relationships described in the Domain Primitives, but the most important ones. The set of networks was divided according to the main concepts being described. It starts with a high level network which describes the general concepts and relationships of the domain knowledge. The followinf figures illustrate the networks.

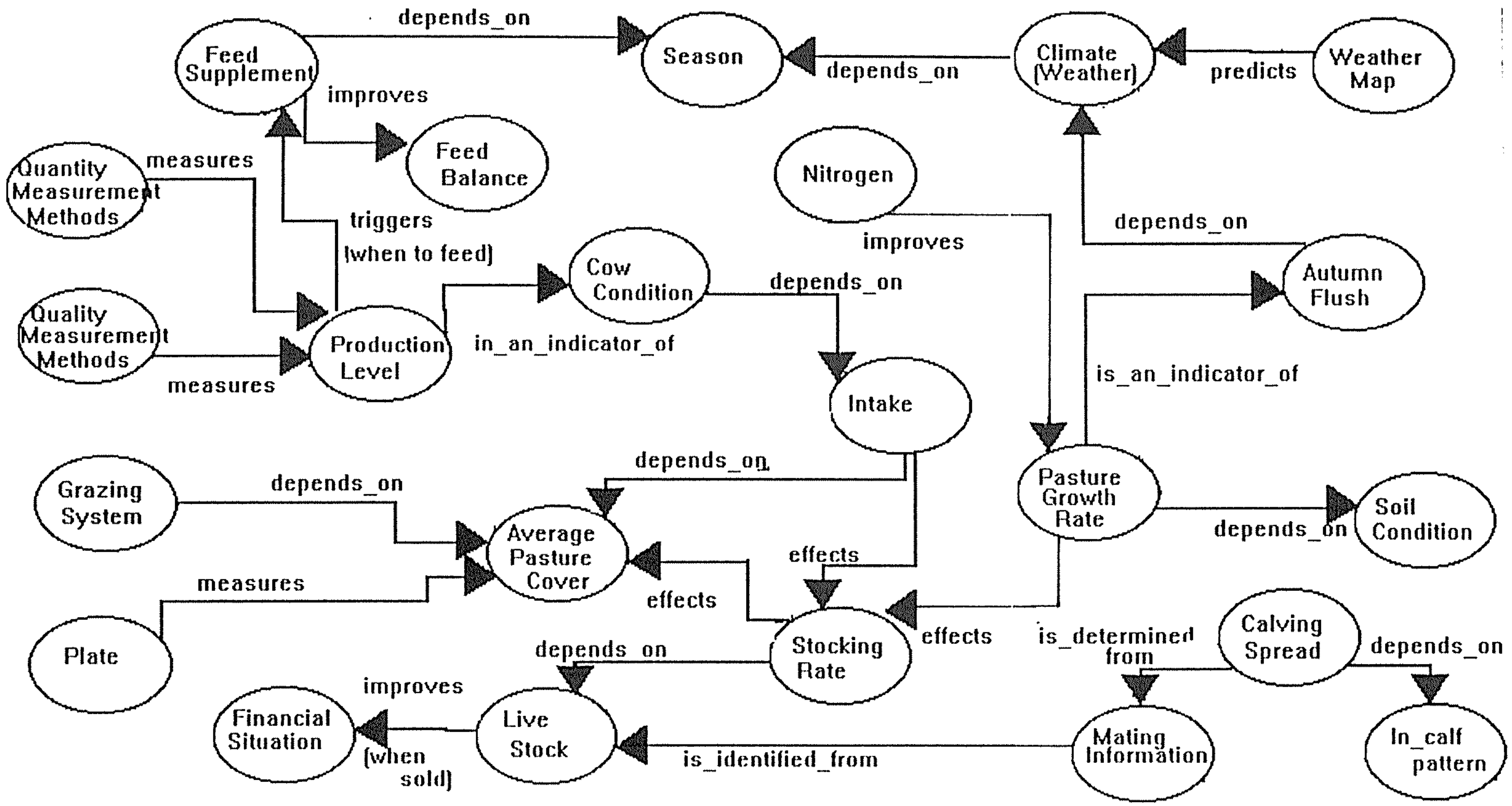


Figure A2.1. High level network

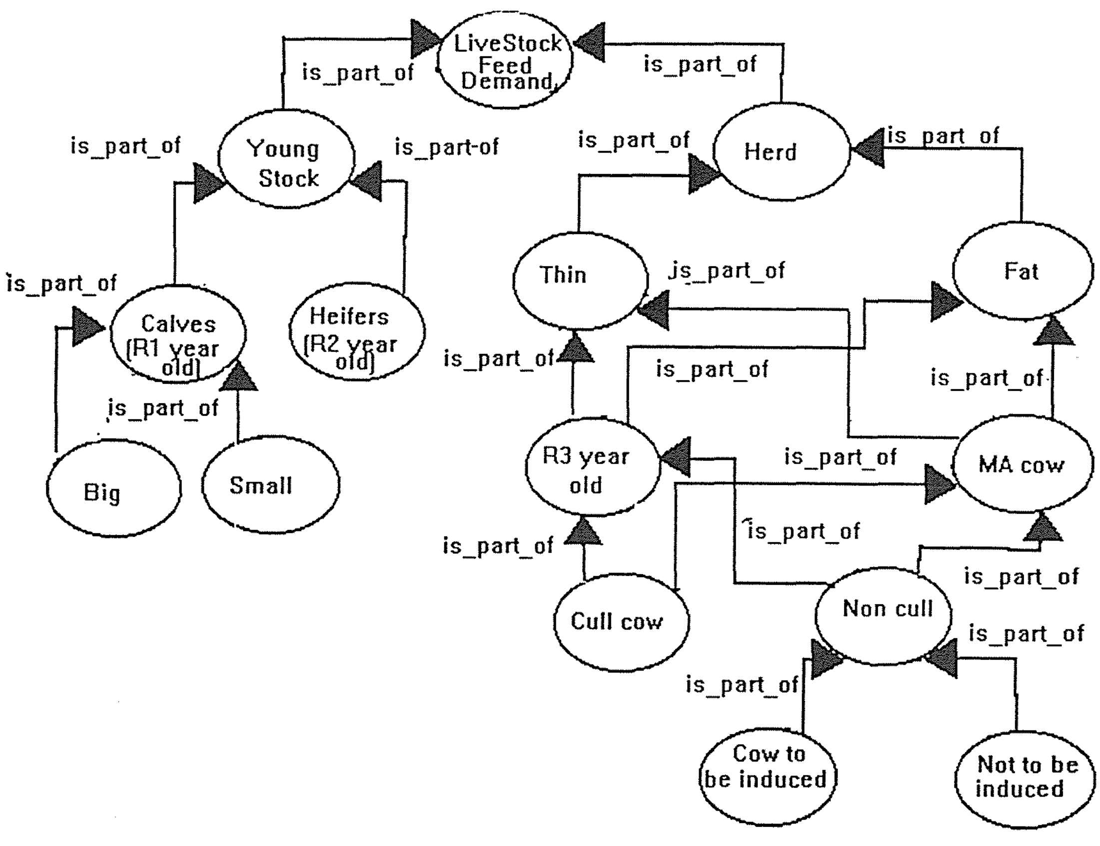


Figure A2.2. LiveStock network - "is_part_of" relations

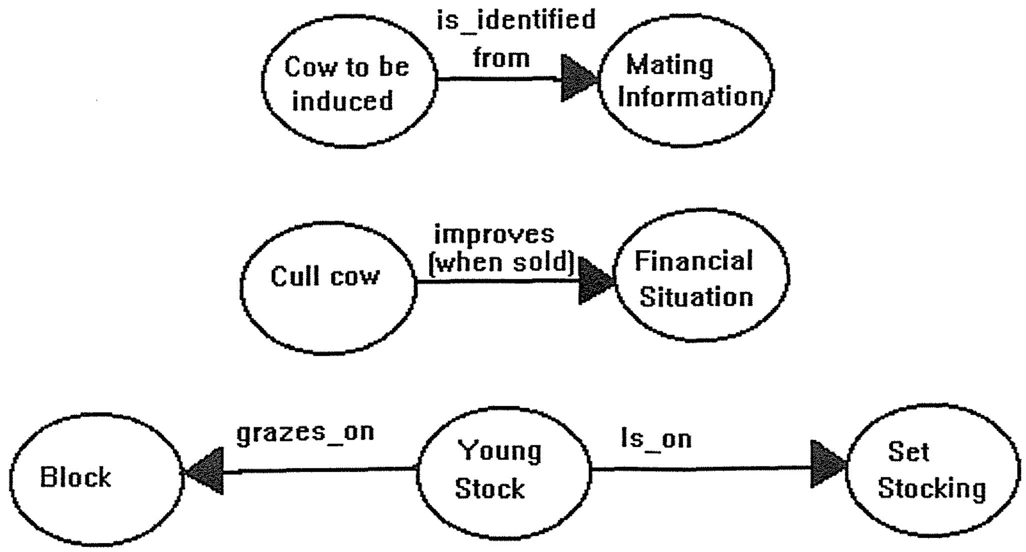


Figure A2.3. Other relations involving livestock

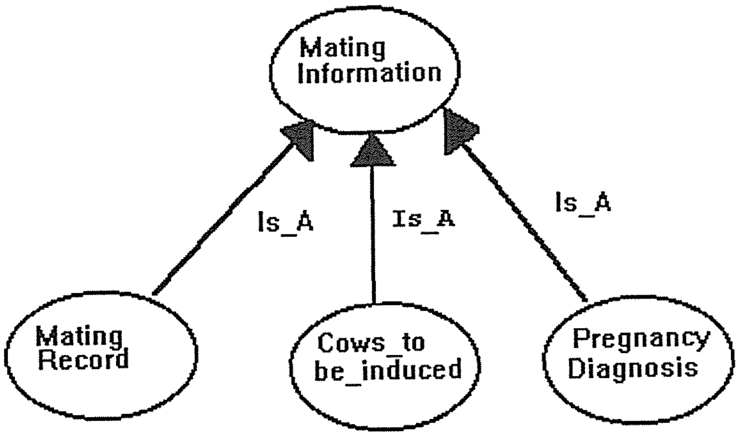


Figure A2.4. Mating information network - "Is_A" relations

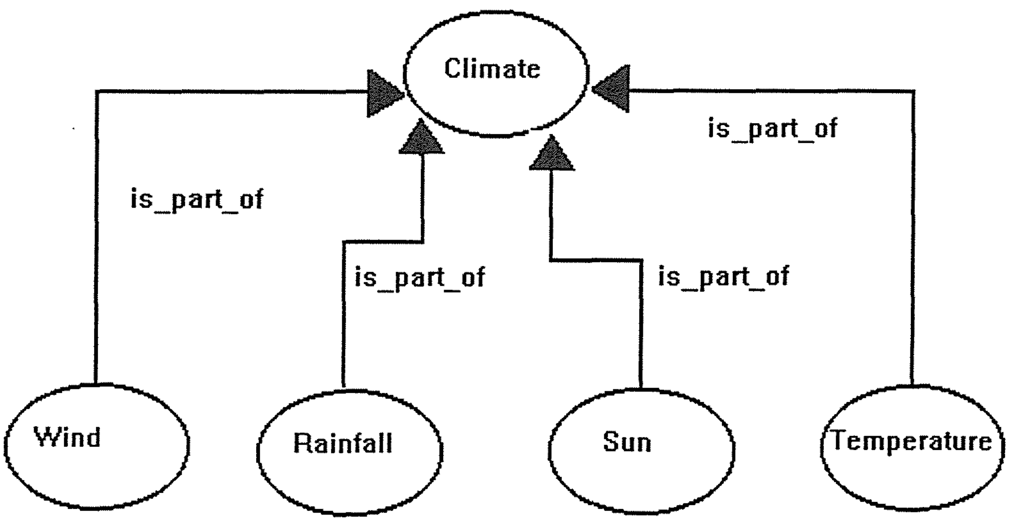


Figure A2.5. Climate network - "is_part_of" relations

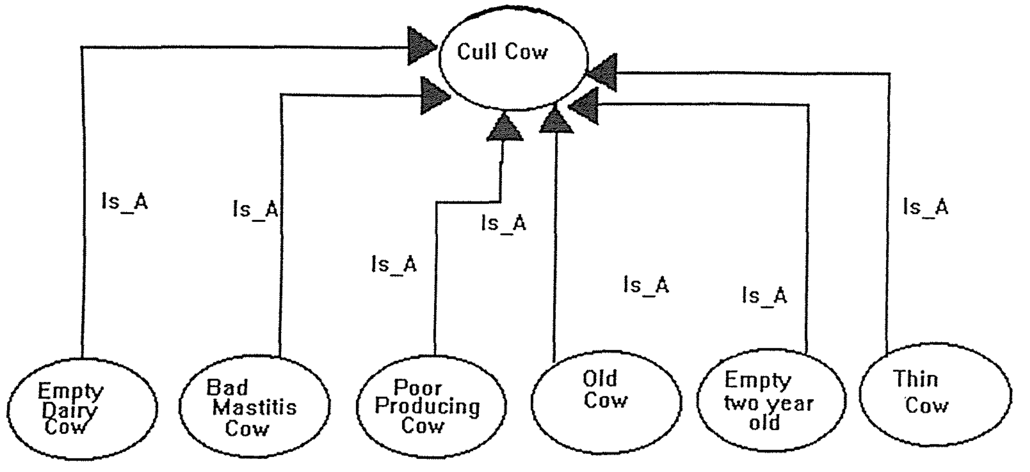
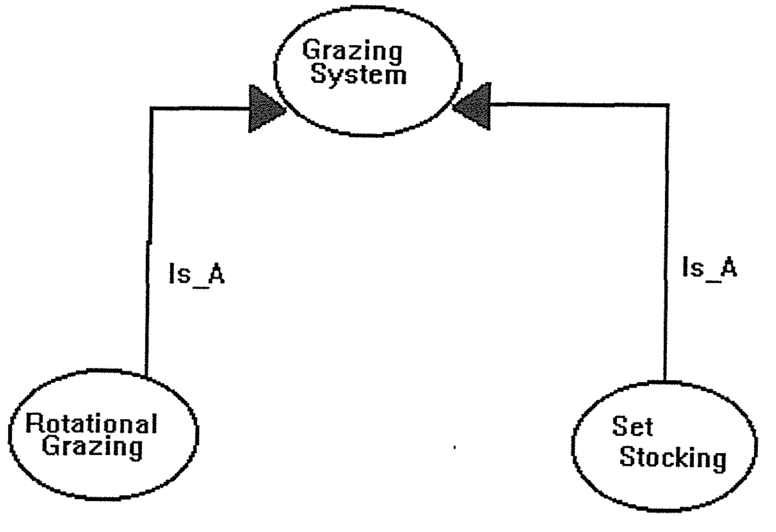


Figure A2.6. Cull cow network - "Is_A" relations

"Is_A" relations



"is_part_of" relations

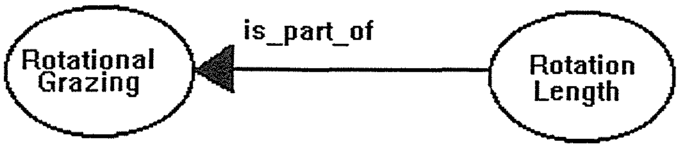


Figure A2.7. Grazing systems network

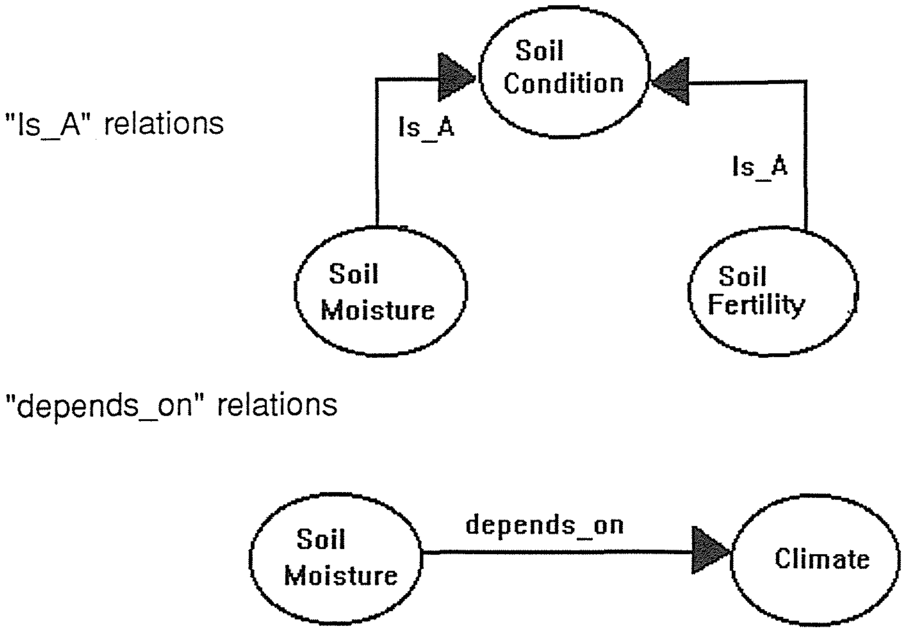


Figure A2.8. Soil condition network

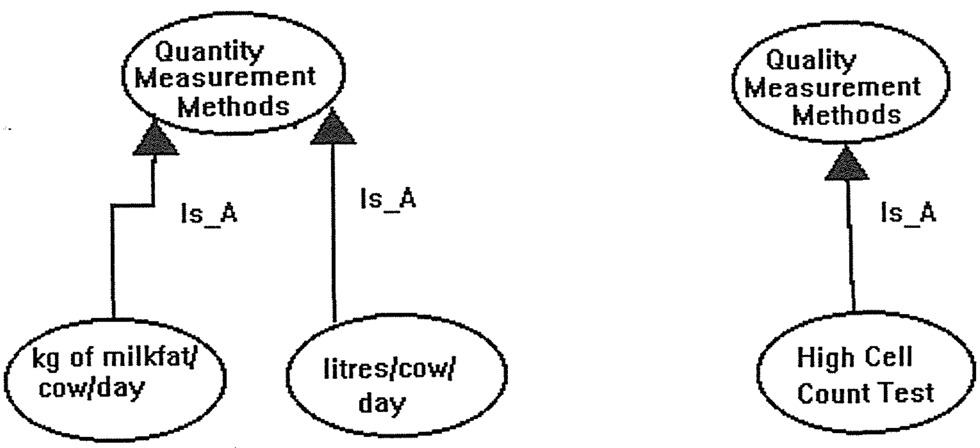


Figure A2.9. Measurement methods network - "Is_A" relations

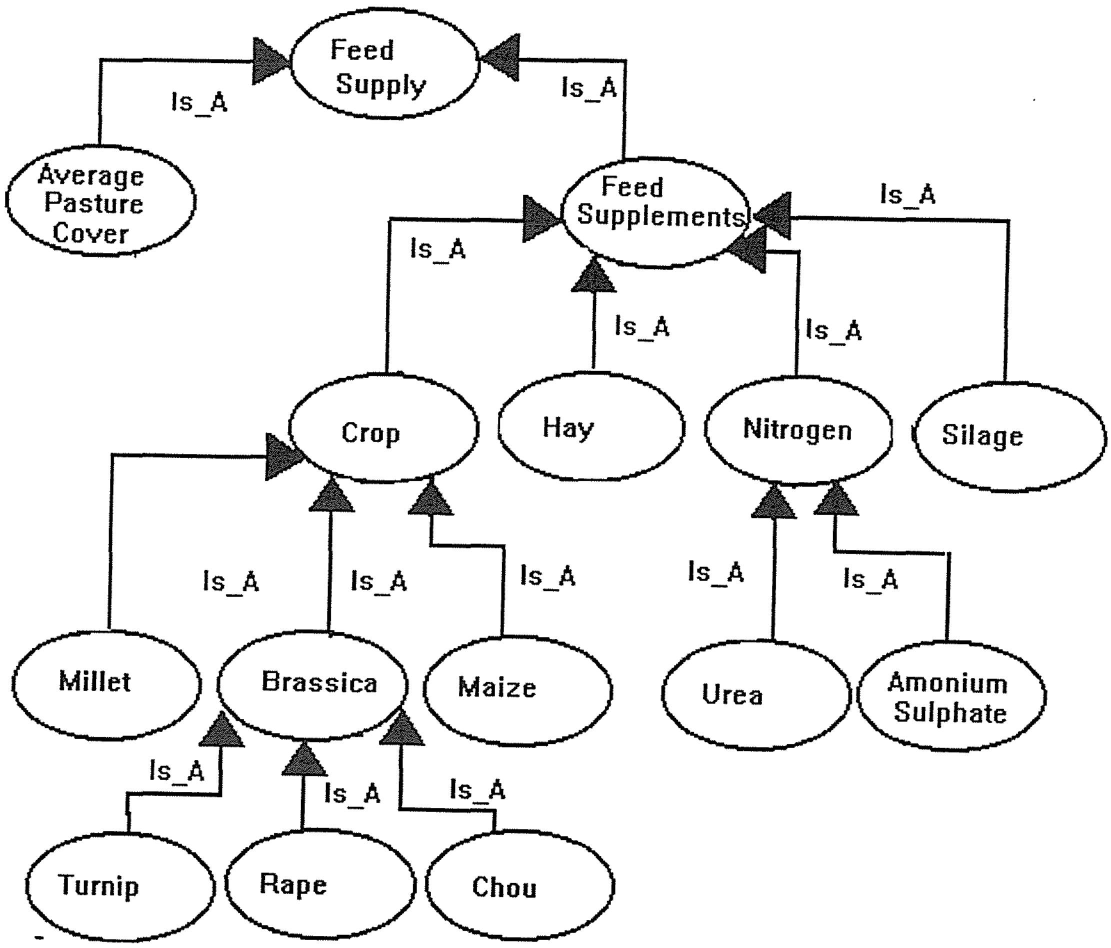


Figure A2.10 Feed supply network - "Is_A" relations

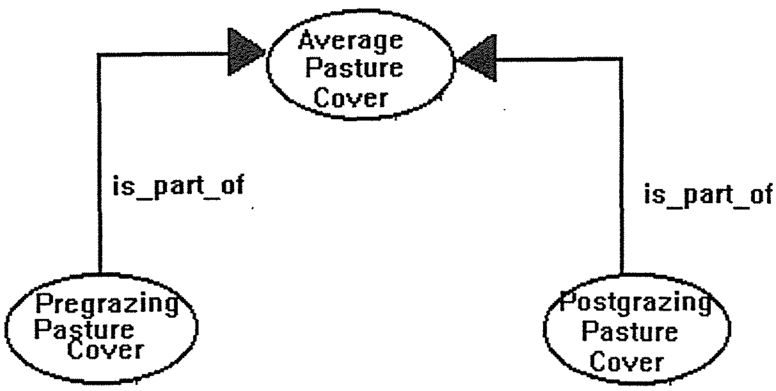


Figure A2.11 Average pasture cover network - "Is_part of" relations

Appendix 3: Inference layer

The inference layer consists of the inference structure of the summer-autumn management domain and the instances of the domain which were instantiated with the primitive inference actions and knowledge roles. Following is the final inference structure designed for the domain (figure A3.1).

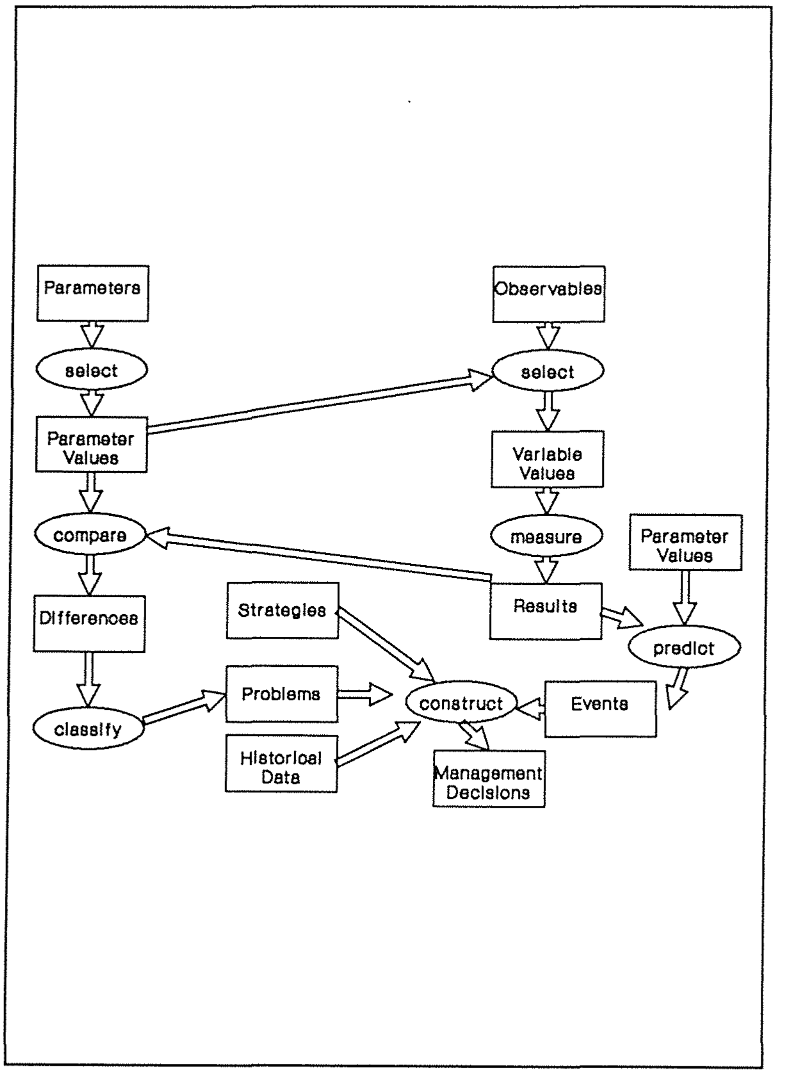


Figure A3.1. Final inference structure of the summer-autumn management period.

A3.1 Primitive inference actions and their input and output knowledge roles

select (Parameters): The inference structure starts with the selection of a number of parameters concerning the structure of the system being developed. The selection of parameters can be compared to an initial study of how an ideal system should work. This Primitive Inference Action also instantiates the parameters to their values.

input: Overall structure of a dairy system.

output: Parameter Values.

select (Observables): Observables are selected from the real system. This selection may be guided by the parameter variables. The observables are transformed into variables through data abstraction.

input: General data concerning the farm current situation.

output: Variable Values.

measure: Before comparing the variables obtained from the system it is necessary to evaluate them. These Primitive Inference Actions transform the observables into results that can be compared to the ideal parameters.

input: Variable Values.

output: Results.

compare: The desired (parameter) values are compared with the result of the test and measure of the actual values. This comparison may generate no difference between the desired and actual values. But, in case a difference is produced, the problem causing the difference is identified.

input: Results and Parameter Values.

output: Differences

classify: This Primitive Inference Action classifies the difference to determine whether it is significant or not.

input: Differences.

output: Problems.

construct: Identifies the actions that have to be taken in order to overcome the problems and maintain the right results.

input: Problems, Strategies, Historical Data and Events.

output: Management Decisions.

predict: It uses some of the results to determine what will happen next in terms of events.

input: Results and Parameter Values.

output: Events.

A3.2 Concepts of the domain which were instantiated with the knowledge roles:

Knowledge Roles	Concepts of the Domain
Parameter (ideal) and Observables (actual)	-Dairy Cow Breed -Types of Supplements -Grazing System Management -Reproductive Performance -Measurement Methods -Soil Types
Parameter and Observable Values	-Rainfall -Cow Condition -Feed Demand (Dry Matter/hectare) -Feed Supply (Dry Matter/hectare) -Type of Grazing System -Animal Intake (kg of dry matter/day) -Pasture Cover (kg of dry matter/hectare) -Pasture Growth Rate -Postgrazing Pasture Cover (kg of dry matter/hectare) -Pregrazing Pasture Cover (kg of dry matter/hectare) -Production Level (kg of milkfat/cow/day) -Reproductive Performance (number of empty cows, conception rate) -Stocking Rate

Knowledge Roles	Concepts of the Domain
Results	-Average cow condition of 4.5 - Pregrazing Pasture Cover of 1800 kg of dry matter/hectare -Conception Rate of 75% -10 empty cows -Stocking rate of 3 cows/hectare -Rainfall of 3 mm overnight

Differences (Examples)	Desired Value	Real Value	Difference
Average Cow Condition	5.0	4.0	1.0
Animal Intake	10.0	8.0	2.0
Milkfat Production	0.7	0.6	0.1
Average Pasture Cover	2000	1800	200

Knowledge Roles	Concepts of the Domain
Problems	-Feed Shortage -Young Stock below target weight -Low in calf rate -Low Production Level -Empty Cows -Low Condition Score -Non ideal weather conditions

Knowledge Roles	Concepts of the Domain
Management Decisions and Strategies (from Planning)	-Feed Silage -Change Rotation Length -Feed Crop -Reduce Cow Intake -Sell unwanted cows -Dry off thin cows -Dry off young cows -Milk the herd once a day -Dry off the herd
Events (from Prediction)	-Make hay and silage -Feed Crop -Identify unwanted cows

A3.3 Relationships of the domain instantiated with the primitive inference actions:

select (Parameters):

- Study the general picture of a dairy farm.
- From this study select the parameters with the ideal values to compare to the real situation on the farm. For the dairy farmers, these parameters can be records of the last years level of production, stocking rate, average pasture cover, drying off date etc.

select (Observables):

- Walk the farm (usually on a fortnightly basis) to analyze the situation and obtain the variables to be tested and measured.

measure:

- Do the pregnancy test to identify empty cows (pregnancy test).
- Condition Score the cows to make sure they are at an acceptable condition.
- Test the milk for milkfat constituents (Herd Test).
- Use the plate meter to measure pasture cover.
- Use levels of milkfat and protein to measure pasture quality.
- Measure how much feed the cows are eating (cow intake) to make sure the production level is being maintained.

compare:

- There are no examples of this primitive inference action as it consists as the comparison of the results of the tests and measurements with the results of the last years. Sometimes these comparison is done based on tacit knowledge.

classify:

- Feed deficiency makes the milk production drop.
- Feed deficiency causes low in calf rate.
- Low in calf rate causes empty cows.
- Young stock below target weight jeopardises the following seasons.
- Condition Scores below acceptable levels cause drop in production and low in calf rates.

construct:

- Rotation length is planned to maximise pasture growth.
- The feedbudget is planned according to feed demand and feed supply.
- The pasture cover management plan consists of not allowing the cover to drop too much.
- The drying off date is planned according to cow condition and pasture cover.

predict:

- The weather predicts when to make hay.
- Residual Pasture Cover predicts production level.
- Production Levels predict cow condition.

Appendix 4: Task layer

The task layer consists of the decomposition of the generic task and was based on the instantiation of the "Manage_the_farm" template (figure A4.1) to the inference layer.

task **Manage_the_farm**

goal: Maximise current milkfat production without jeopardising next season's yield.

task structure

select(Observable--->Variable Values)

select(Parameters--->Parameter Values)

measure(Variable Values---> Results)

monitor(Results and Parameter Values--->Differences)

diagnose(Differences--->Problems)

plan(Problems and Strategies--->Management Decisions)

predict(Results and Parameter Values--->Events)

Figure A4.1. Generic task template of the summer-autumn management period.

This instantiation detected the necessity of decomposing the generic task into subtasks. Following the four subtasks which resulted from the decomposition of the generic task "Manage_the_farm".

task **Planning Management Decisions**

goal: Organize a plan of actions that have to be taken in order to maintain the production level of current system without risking next season's production.

task structure

construct(Problems,Strategies, Historical Data and Events----> Management Decisions)

The second activity consists of assessing the events on the farm:

task **Monitoring Events**

goal

Make sure the results being obtained are according to the goals to be achieved.

task structure

select(Observable--->Variable Values)

select(Parameters--->Parameter Values)

measure(Variable Values--->Results)

compare(Results and Parameter Values--->Differences)

The third activity consists of identifying the cause of problems:

task **Diagnose Problems**

goal: Identify cause of problems

task structure

select(Observable--->Variable Values)

select(Parameters--->Parameter Values)

measure(Variable Values--->Results)

compare(Results and Parameter Values--->Differences)

classify(Differences--->Problems)

The fourth activity consists of the prediction of events on the farm:

task **Prediction of Events**

goal: Identify the future state of the events on the farm in order to take specific actions to avoid problems.

task structure

select(Observable--->Variable Values)

select(Parameters--->Parameter Values)

measure(Variable Values--->Results)

predict(Results and Parameter Values--->Events)

A4.1 Task decomposition

This is the main part of the task layer, the specification of the generic tasks: Planning Management Decisions, Monitoring Events, Diagnose of Problems and Prediction of Events. It consists of the decomposition of the generic tasks into the subtasks and goals associated with them. This specification is the overall framework of the tasks performed by the dairy farmers in order to achieve their goals. The tasks described are related to the inference structure of figure A3.1.

For all subtasks:

- The names of the primitive inference actions (are italicised in the task structure.

- The arrows in the task structure describe the relation between input and output knowledge roles which are represented by concepts of the domain.

- In case a Transfer Task is necessary for the achievement of the subtask being described, its name will be italicised.

<u>Subtasks of the generic task: Monitoring Events</u>

subtask: Measure Pasture Cover

goals: - Make sure there will not be a feed deficit on the next days.

- Verify if the cows are being well fed at the moment.

task structure:

select(Pasture Cover--->Average Pasture Cover)

select(Pasture Cover System--->Ideal Pasture Cover)

Walk the farm

Select a paddock as a sample

measure(Pasture Cover--->Results of Pasture Cover Measurement)

If the month is January, February or March

Then

Calculate Pasture Cover informally once a fortnight;

Else

Measure Pasture Cover using the Plate Meter once a week

End If

compare(Ideal Pasture Cover with Results of Pasture Cover Measurement---> Differences in the Pasture cover)

subtask: Measure Production Level

goals:- Confirm that the animals are being well fed.

- Make sure the production goals are being achieved.

task structure:

Repeat every day for herd

select(Production Level--->Production/cow/day)

select(Production System--->Ideal Production)

Measure total of litres

Measure kg of milkfat

Divide total of litres by number of cows

Divide total of milkfat by number of cows

measure(Milkfat/cow/day--->Results of milkfat/cow/day)

measure(Litres/cow/day--->Results of litres/cow/day)

compare(Ideal Production with Results of milkfat/cow/day ---> Differences in Production Level)

compare(Ideal Production with Results of litres/cow/day ---> Differences in Production Level)

End.

subtask: Measure Body Condition (Condition Score)

goals:- Confirm that the cows are being well fed.

- Identify thin cows to be dried off.

task structure:

select(Condition Score--->Condition Score of the cows)

select(Condition Score System--->Ideal Condition Score)

For herd

measure(Condition Score of the cows--->Results of Cow Condition)

compare(Results of Cow Condition with Ideal Condition Score ---> Differences in Condition Score)

subtask: Monitor Crop

goal: - Make sure the quality of the crop is not being compromised.

task structure:

select(Type of Crop--->Crop)

select(Supplement System--->Ideal Quality)

measure(Crop--->Results)

compare(Crop with Ideal Quality--->Differences in Quality)

subtask: Do the Pregnancy test

goal: Identify empty cows.

task structure:

select(Cows to be tested--->Selected Cows)

select(Pregnancy Test System--->Ideal Percentage of Pregnant Cows)

measure(Selected Cows--->Results of Pregnancy Test)

compare(Results of Pregnancy Test with Ideal Percentage of Pregnant Cows--->Differences in Percentage of Pregnant Cows)

<u>Subtasks of the generic task: Diagnose Problem</u>

subtask: Identify poor producing cows

goal: - Maintain production level

task structure

Measure production level

If the cow is producing less than 0.1 kg of milkfat/day

Then

classify(Difference in Production Level--->Poor producing cow)

End If.

subtask: Identify thin cows

goal: Maintain body condition

task structure

Measure Body Condition

If the condition score if <= 3.5 and there is no time to improve body condition

Then

classify(Difference in Condition Score--->Thin Cow)

End If.

subtask: Identify Feed Shortage (Between April and August)

goal: Maintain cows well fed

Make sure that will be enough feed for the cows at calving time and early lactation.

task structure:

Measure Pasture Cover

If Average Pasture Cover is below 1400 kg of Dry Matter/ hectare **and** supplements are not enough to cover the deficit

Then

classify(Difference in Pasture Cover-->Feed Shortage)

End If

If Average Pasture Cover will fall below 2200 kg of Dry Matter/hectare in late July

Then

classify(Difference in Pasture Cover-->Feed Shortage)

End If.

subtask: Identify drop in production level

goal: - Maintain Production Level

task structure

Measure Production Level

If production has fallen below 0.6 kg of milkfat/cow/day

Then

classify(Difference in Production Level--->Drop in Production Level)

End If.

subtask: Identify quality problem of summer crop

goal: Graze crop before its quality decline

task structure:

Measure Crop

If crop has grown too much

Then

classify(Difference in Supplement--->Quality problem with summer crop)

End If.

Subtasks of the generic task: Predict Events

subtask: Predict production level by estimating cow intake.

goal: - Maintain production level by predicting how much the cows are eating.

task structure

For one paddock

Measure Pasture Cover before grazing (Pregrazing)

Measure Pasture Cover after grazing (Postgrazing)

Subtract Postgrazing from Pregrazing to get Residual Pasture Cover

Divide the result of the subtraction by number of cows

If the cow intake is below 10 kg of Dry Matter/day

Then

predict(Results of Residual Pasture Cover--->Improve Animal's Diet)

End If.

Subtasks of the generic task: Plan Management Decisions

subtask: Change Rotation Length

goal: - Maintain cow condition

- Take advantage of high pasture growth rates

task structure

Measure Pasture Cover

If *Identify Feed Shortage* is true

Then

If rotation length can be lengthened

Then

construct(Feed Shortage and Rotation System Strategy--->Lengthen Rotation System)

Else

construct(Feed Shortage and Feed Supplement Strategy--->Feed Supplement).

End If

End If.

subtask: Feed Supplement

goals: - Maintain Production Level

task structure

If *Identify Drop in Production Level* is true

Then

If crop is available

Then

If crop quality is declining

Then

construct(Problem:Drop in Production Level; Crop Strategy---> Feed Crop)

Else

construct(Problem:Drop in Production Level; Silage Strategy---> Feed Silage)

End IF

Else

construct(Problem:Drop in Production Level; Silage Strategy--->Feed Silage)

End If

End If.

subtask: Identify Cull Cows

goal: - Maintain cow intake by selling marked cows

- Improve Average Pasture Cover

task structure

If *Identify feed shortage* is true **and**

Predict empty cows is true

Then

construct(Problem:Empty Cows; Strategy for Empty Cows--->Sell empty cows)

End If

If empty cows have been already sold **and**

Identify feed shortage is true **and**

Identify poor producing cows is true

Then

construct(Problem:Poor Producing Cows; Strategy for Poor Producing

Cows ---> Sell Poor Producing cows)

End If

If empty and poor producing cows have been already sold **and**

Identify feed shortage is true **and**

Identify thin cows is true

Then

construct(Problem:Thin Cows; Strategy for Thin Cows--->Sell Thin cows)

End If.

subtask: Milk cows once a day

goal: Maintain Average Pasture Cover and Cow Condition

task structure

If *Identify Feed Shortage* is true **and**

Cull cows have already been sold **and**

Cow Condition can not be used

Then

construct(Problems:Feed Shortage and Cow Condition; Drying off Strategy--> Milk herd once a day)

End If.

subtask: Dry off the herd

goal: Maintain Average Pasture Cover and Cow Condition

task structure

If *Identify Feed Shortage* is true **and**

Cull cows have already been sold **and**

Cow Condition can not be used **and**

The herd is being milking once a day already

Then

construct(Problems: Feed Shortage and Cow Condition;Drying off Strategy--> Dry off the Herd)

End If.