The Development of a Fuzzy Decision-Support System for Dairy Cattle Culling Decisions

by Mark Strasser
Department of Agricultural and Biosystems Engineering
McGill University
May 1997

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements of the degree of Masters of Science

0-612-29794-2

Table of Contents

Abstract

In order to investigate the use of fuzzy logic in decision-support systems (DSS) for dairy cattle breeding, a first-generation prototype software system was developed. The objectives were to determine the advantages and limitations of fuzzy logic for this type of application, and to establish a basis for the development of more complete DSS in the future. The goal of the prototype DSS was to make culling decisions on the basis of monthly production data. An analysis of the development process of this prototype demonstrated the importance of creating a thorough methodology for the elicitation and implementation of knowledge. A framework for the development of fuzzy decision-support systems was established, consisting of four phases: the project groundwork phase, elicitation of knowledge from the expert, implementation of that knowledge, and system validation. In this framework, it is proposed that, in the case of multiple experts, knowledge can be amalgamated or aggregated. Once this framework was established, a second-generation prototype DSS was developed. Contrary to the first-generation prototype, where the encoded expertise was limited to three experts from the same domain, the second-generation prototype considered the knowledge of two individuals from each of three domains (Dairy researchers, Producers, and Dairy herd improvement specialists). An aggregation approach was used which involved the development and maintenance of separate modules, each containing the compiled expertise of one of the six experts.

Résumé

Un prototype de logiciel de première génération a été développé dans le but d'étudier la possibilité d'utiliser la logique floue dans des systèmes d'aide à la décision (SAD) pour l'élevage de la vache laitière. Les objectifs étaient de déterminer les avantages et les limites de la logique floue pour ce type d'utilisation, et d'établir les bases nécessaires pour le développement de SAD plus complets. Le but du prototype de SAD était d'amener à prendre la décision de réforme des vaches en fonction des données sur leur production mensuelle. L'analyse du processus de développement de ce prototype a démontré l'importance de créer une méthodologie consciencieuse pour l'obtention et l'implantation de connaissances. Une structure de base pour le développement de systèmes d'aide à la décision utilisant la logique floue a été établie. Elle consiste en quatre phases: une phase préliminaire, une phase d'acquisition des connaissances de l'expert, une phase d'implantation de ces connaissances et une phase de validation du système. Dans cette structure, il est proposé que, dans le cas d'experts multiples, les connaissances peuvent être amalgamées ou agrégées. Une fois que cette structure fut établie, un prototype de SAD de deuxième génération a été développé. Contrairement au prototype de première génération où l'expertise codée se limitait à trois experts dans le même domaine, le prototype de deuxième génération prenait en compte les connaissances de deux individus dans chacun des trois domaines (Chercheurs, Producteurs et Spécialistes en amélioration du troupeau). Une approche en agrégation qui implique le développement et l'entretien de modules séparés a été utilisée, chacun contenant une compilation de l'expertise d'un seul des six experts.

Acknowledgements

I would like to extend my gratitude to a number of individuals. Firstly, a great many thanks are extended to my two supervisors, Dr. Kevin Wade, and Dr. Robert Kok and to the other members of my thesis committee, Dr. René Lacroix, and Dr. Flann Hayes, for their support and friendship.

Further thanks are extended to the experts who graciously provided their time and knowledge during the course of this project: Denis Marchand, Chantal Cliche, Steeve Adam, Dr. Humberto Monardes, Chris and Mike Gasser, and Hans Kaiser.

I am also very grateful to my past and present colleagues of the Information Technology Group in Animal Science for their friendship and support: Xingzhu Yang, Ainsley Archer, and Diederik Pietersma.

Much thanks are extended to the Department of Animal Science support staff. In particular I would like to thank Barbara Stewart, and Sharon O'Toole, for making the impossible look easy, as well as, Lisa Gloutney, and Joyce Reid, who have since moved on to "greener pastures".

A very special thanks are extended to my family who were a constant source of support.

This project was made possible with the financial support of the Natural Sciences and Engineering Research Council of Canada and the Cattle Breeding Research Council of Canada. Data used in Chapter 1 were provided by the Québec Dairy Herd Analysis Service (Programme d'analyse des troupeaux laitiers du Québec). Data used in Chapter 3 were provided by the Macdonald Campus animal research facility.

This thesis is dedicated to the memories of my grandmother, Rose Brenner, and to my friend and colleague, Ken Foran.

Preface

The following paragraphs are included as part of the requirements as stipulated by the Faculty of Graduate Studies and Research:

"Candidates have the option of including, as part of the thesis, the text of one or more papers submitted or to be submitted for publication, or the clearly-duplicated text of one or more published papers. These texts must be bound as an integral part of the thesis.

If this option is chosen, CONNECTING TEXTS THAT PROVIDE LOGICAL BRIDGES BETWEEN THE DIFFERENT PAPERS ARE MANDATORY. The thesis must be written in such a way that it is more than a mere collection of manuscripts; in other words, results of a series of papers must be integrated.

The thesis must still conform to all other requirements of the "Guidelines for Thesis Preparation". THE THESIS MUST INCLUDE: A Table of Contents, an abstract in English and French, an introduction which clearly states the rationale and objectives of the study, a comprehensive review of the literature, a final conclusion and summary, and a thorough bibliography or reference list.

Additional material must be provided where appropriate (e.g. in appendices) and in sufficient detail to allow a clear and precise judgement to be made of the importance and originality of the research reported in the thesis.

In the case of manuscripts co-authored by the candidate and others, THE CANDIDATE IS REQUIRED TO MAKE AN EXPLICIT STATEMENT IN THE THESIS AS TO WHO CONTRIBUTED TO SUCH WORK

AND TO WHAT EXTENT. Supervisors must attest to the accuracy of such statements at the doctoral oral defence. Since the task of the examiners is made more difficult in these cases, it is in the candidate's interest to make perfectly clear the responsibilities of all the authors of the co-authored papers."

This thesis contains a total of three manuscripts, two of which have been submitted for publication. Chapter 1 has been submitted to the Journal of Dairy Science. Chapter 2 has been submitted to the Journal of the Canadian Society of Agricultural Engineering.

The following is a breakdown of the contributions made by the authors towards the preparation of the three manuscripts submitted as part of this thesis:

<u>CHAPTER I</u>

The candidate was responsible for development of the decision support system described. The candidate also conducted the experiments used in an analysis of the system's performance. Dr. René Lacroix was responsible for the preparation of the manuscript and the analysis of the system. Assistance was provided by Dr. Robert Kok, and Dr. Kevin Wade through their general guidance and editorial input in the preparation of the manuscript.

<u>CHAPTER II</u>

The candidate was responsible for conducting the research and preparing the manuscript. Assistance was provided by Dr. René Lacroix, Dr. Robert Kok, and Dr. Kevin Wade through their general guidance and editorial input in the preparation of the manuscript.

<u>CHAPTER III</u>

The candidate was responsible for conducting the research and preparing the manuscript. Assistance was provided by Dr. René Lacroix, Dr. Robert Kok, and Dr. Kevin Wade through their general guidance and editorial input in the preparation of the manuscript.

List of Tables

List of Figures

List of Appendices

INTRODUCTION AND LITERATURE REVIEW

I. Introduction

Dairy farming is a business in which profitability is the main objective. Profitability is directly related to the everyday expenses and income of a farm operation (Mainland, 1994). Almost all dairy farmers try to increase their income by increasing milk yield, which also serves the dual purpose of diluting maintenance and other fixed costs (McDaniel, 1994). This is often accomplished through voluntary culling, which is the removal of lower producing cows and their replacement with higher producing ones. Unfortunately, many cows are culled involuntarily. Involuntary culling occurs when the cow has major physical problems, for example, the inability to conceive, mastitis, poor disposition, or udder and health problems (Rogers *et al.*, 1988). When involuntary culling must be performed, this not only results in a decrease in revenue, due to the removal of a cow during the course of a lactation, but there are also other factors which affect profitability including increased veterinary costs, disposal costs, and a lower carcass price at the slaughterhouse (Monardes, 1992). Any reduction in involuntary culling rates are of benefit to the dairy producer for three reasons: the diminished probability of having to cull high producing cows, more opportunities to cull for voluntary reasons (to increase production), and a decrease in rearing costs (Rogers *et al.*, 1988).

The economic benefits of improving culling policies have been proven. The "top 20%" of dairy producers have significantly lower replacement costs compared to the "bottom 20%" (Pellerin, 1993). Therefore, if it is possible to acquire expertise from the "top" producers, or other experts, and make it available for use on less productive farms, this would potentially increase the latter's profitability. Traditionally, consultants were expected to provide the producer with the information necessary for herd improvement. Over time, agencies have surfaced offering more sophisticated services to the dairy producer, for example, dairy herd improvement agencies (DHI). Through such tools as management reports, these agencies have attempted to promote an increase in profitability by making available a greater amount of information on which to base a decision,. Unfortunately, dairy producers have stated that they do not consult their DHI records due to a lack of time and a lack of comprehension (Pellerin *et al.*, 1994, Smith, 1989). Often a consultant must concurrently be made available to explain the recommendations to the producers.

This exposes the problem of expert availability. Human experts are costly and in short supply (Levins and Varner, 1987). Furthermore, Willett and Andrews (1996) express concern that the various budget cuts may also affect the availability and quality of experts. It is, therefore, suggested that these last problems can be alleviated with the development of computer systems which could be used to mimic the reasoning processes of experts (Willett and Andrews, 1996, Levins and Varner, 1987). Expert knowledge can be captured for use in the artificial intelligence tools, such as expert systems (ES), which can include knowledge from a variety of sources, including, human experts, research results and government policy (Greer *et al.*, 1994). However, expert systems have not attained their forecasted level of acceptance. It has been suggested that this is due to their inability to deal with the vagueness and uncertainty associated with the human reasoning process. Fuzzy logic can be used overcome these shortcomings (Lacroix *et al.*, 1994).

II. Literature Review

Culling

Dairy producers have two main motives for culling their animals; culling for voluntary reasons (eg. low production) and culling for involuntary reasons (eg. breeding problems) (Pellerin, 1993; Monardes, 1992; Rogers *et al.* 1988; Dentine *et al.*, 1987). Other reasons for involuntary culling include: mastitis, poor disposition, udder and health problems, and foot and limb problems (Monardes, 1992; Rogers *et al.*, 1988; Young *et al.*, 1983). In Quebec, the primary reason for culling dairy cattle is low production, the second being reproductive problems (Monardes, 1992). A study conducted by the USDA also showed that most grade cows leave the herd due to low production (Dentine *et al.*, 1987). This can be contrasted with culling practices in East Anglia, England, where dairy cows were culled primarily because of their failure to conceive, followed by low production (Young *et al.*, 1983).

Many models have been developed for the optimization of culling decisions. Of these, many use a dynamic programming approach (Kennedy and Stott, 1993; Rogers *et al.*, 1988; Stewart *et al.*, 1978; Stewart *et al.*, 1977). One study employed the use of a Markovian model (Ben-Ari *et al.*, 1983). By examining the variables considered by these models it is possible to get an understanding of the factors influencing the culling decision. Variables considered include: age, milk yield, milk fat percent, reproductive status, health state, body weight, reasons for disposal, lactation number, stage of lactation, probabilities of involuntary removal, ability to conceive, genetic improvement, variation in production, repeatability of production, death and survival to the next lactation, revenue from milk, calves, culled cows; costs associated with feeding, health, replacement, housing, cow depreciation, interest; and finally, availability of a replacement cow is also considered (Kennedy and Stott, 1993; Rogers *et al.*, 1988; Van Arendonk, 1988; Ben-Ari *et al.*, 1983; Stewart *et al.*, 1978; Stewart *et al.*, 1977).

Many authors stress the importance of decreasing the incidence of involuntary culling. Lowering involuntary culling rates allows the farmer to cull more often for reasons of low production, thus improving milk yields and profitability (Pellerin, 1993; Rogers et al, 1988). By extension, culling for involuntary reasons does not ensure that there exists a better performing replacement cow and there may be other costs associated with the culled animal (eg. veterinary costs, transportation costs, and decreased value at the slaughterhouse) (Monardes, 1992).

The costs associated with culling are significant. The cost of replacement alone (not considering lost revenue) for the average Québec dairy farm approaches 25% of the total costs related to milk production (Pellerin, 1993). The timing of the culling decision is important in that the farmer must, during the animal's productive years, be able to recuperate the cost of rearing the animal (Dentine *et al.*, 1987). This decision must consider that a cow only begins to show a profit in the third lactation (Pellerin, 1993).

In Québec, there has been a sharp rise in the incidence of involuntary culling (Monardes, 1992; Pellerin, 1993). During the period of Jan. 1980 to Dec. 1989, culling for involuntary reasons rose from 47% to 72% of all cows culled, while culling for voluntary reasons decreased from 53% to 28% (Monardes, 1992). A 1988 study showed that a reduction of involuntary culling rates of 2.9% resulted in an increase of $22 U.S. in net revenue per cow per year (Rogers *et al.*, 1988). A study conducted in Québec determined that each dairy producer spends approximately $32 709 to replace an average of 15 cows per year. For the whole province, this represents a cost of $442 million per year. The costs associated with a "top 20%" producer are much lower than those of a "bottom 20%" producer. A "top 20%" producer pays $22 900 to replace 15 cows while the "bottom 20%" producer pays up to $49 200 (Pellerin, 1993). Decreasing the costs associated with involuntary culling combined with an increase in herd life results in a considerable increase in revenue (Pellerin, 1993). From this it is evident that any improvement in culling practices will influence farm profitability.

Decision-Support Systems

Decision support systems (DSS) are computer programs that contain the encoded knowledge of experts. This knowledge is contained in applications such as simulation models, expert systems, databases, or spreadsheets (Greer *et al.*, 1994). DSS are designed to automate data analysis and to perform complex decision-making through the use of human-like reasoning (Allore *et al.*, 1995). DSS can best be summarized as being a tool used "to increase the decision-making power of the human by providing easy access to *useful* data, information and knowledge" (Rauscher, 1995).

The architecture of DSS exposes the user to the decision-making process and the reasoning behind a decision. From the developers perspective, since knowledge is encoded in a straightforward fashion, it more easily allows for the discovery of weaknesses in the knowledge base and to better pinpoint areas where further research is needed (Allore *et al.*, 1995). Another spinoff from DSS development is that it brings researchers and producers together which increases the effectiveness of research and development (Cox, 1996).

There exist some major considerations when developing DSS. DSS must be flexible and easily modifiable to account for any changes in domain policy, rules, limitations, and opportunities. Also, the more integrated a DSS becomes the more difficult and expensive it is to construct (Laacke, 1995). Another consideration is the design of the user interface. Cox (1996) expresses concern that DSS are often constructed with an interface which is too complex to be operated effectively by farmers who have limited experience with computers. Cox explains this by asserting that "information is not knowledge, particularly when the information is presented in a way that is confusing, embodies a conflicting set of values, and fails to recognize the decision-making style of the client group (p. 361)". It is therefore extremely important to consider the end-user when creating DSS.

Many DSS have been constructed for a variety of applications. Recent examples include a GIS-based DSS for the management of livestock production systems (Jain *et al.*, 1995), a DSS for reducing pesticide use (Secher *et al.*, 1995), and a DSS for ecosystem management (Laacke, 1995). In examining research performed within the dairy sector it can be seen that almost all DSS use expert systems. Examples of these systems will be described in the following section.

Expert Systems

One of the most commonly used tools for developing decision-support systems is the expert system (ES). Expert systems typically consist of two main components, the knowledge base and the inference mechanism (McKinion and Lemmon, 1985). The knowledge base consists of the factual knowledge, the procedural rules, the assumptions, and the heuristics that an expert uses in performing a particular task (Greer *et al.*, 1994). Procedural or production rules are the most commonly used method for representing knowledge in the knowledge base

(McKinion and Lemmon, 1985) and are commonly constructed using an IF...THEN syntax. The inference mechanism is the expert system component which is used to interpret the knowledge encoded in the knowledge base. It determines how the knowledge relates to other knowledge and matches this knowledge with relevant data (Schneider *et al.*, 1996). One of the differences between the knowledge base and the inference mechanism is that whereas the knowledge base can expand as new knowledge becomes available, the inference mechanism remains static (McKinion and Lemmon, 1985).

The inference mechanism is divided into two parts, as described by Schneider *et al.*, (1996). The first part is a "blackboard" which is used to store the data and the intermediate results from the inference procedure. The second part is the inference engine which contains the procedures of the inference process based upon the *modus ponens* rule. Whenever the premise (the IF part) of a rule is true, the rule fires, and the conclusion (or the THEN part) of the rule is also true. An example rule is: IF production is very low THEN cull the cow. There are three different inference procedures: forward chaining, backward chaining and direct chaining. According to Spahr *et al.* (1988), forward chaining is most appropriate for problems where many goals are involved and there is a limited amount of data. In this case the inference is data driven; the mechanism starts with the observations and tries to reach a conclusion. They describe backward chaining as being goal driven; the mechanism commences with a hypothesis and searches for data to support this hypothesis. Direct chaining, the third method for inference, utilizes relational lists to perform the inference (Schneider *et al.*, 1996).

Expert systems evolve over a number of stages. The first stage involves the creation of a prototype, allowing the developers to gain some insight into the complexity of the subject domain (Smith, 1989). The prototype can also be used as a demonstration model to garner financial support for a larger project. The prototype can then be expanded into a more detailed and precise model of the subject area. The four steps involved in the creation of an ES knowledge base are problem identification, knowledge base formalization, testing and prototype revision, and implementation (Plant and Stone, 1991). Problem identification is used to determine whether the expert system format is appropriate for the subject area which is to be modelled. For example the type of knowledge must be examined to determine whether the knowledge is procedural or declarative. Procedural knowledge is not appropriate for expert systems, but classical procedural programming is more appropriate to represent it. Knowledge base formalization is the process of acquiring and encoding the knowledge of the experts. Testing and prototype revision includes the verification and validation of the prototype and its revision as necessary. Finally, implementation, is the development and distribution of the final product.

Expert systems can be applied to a number of different scenarios. Harrison (1991, p.267) stated that expert systems can be used:

> "(1) As a stand-alone advisory system for the specific knowledge domain, perhaps with monitoring by a human expert
> (2) to provide decision-support for a high-level human expert
> (3) to allow a high-level expert to be replaced by a subordinate expert, aided by the expert system
> (4) as a delivery system for extension information
> (5) to provide management education for decision makers
> (6) for dissemination of up-to-date scientific information, in a readily accessible form, to agricultural researchers and advisers."

Over the past decade many expert systems have been developed for implementation in the agricultural sector. For example, expert systems have been developed for monitoring the production of laying hens (Lokhorst and Lamaker, 1996), for crop planning (Nevo *et al.*, 1994), for the diagnosis of potato diseases (Boyd and Sun, 1994), for grain marketing analysis (Thieme *et al.*, 1987), and for the identification of honey bee diseases, parasites, pests, and predators (McClure *et al.*, 1993). There have also been a number of expert systems developed for dairy cattle management, including some for the evaluation of reproductive performance and management (Domecq *et al.*, 1991), for overall dairy herd management (Pellerin *et al.*, 1994), and for reproductive problems (Levins and Varner, 1987). Dairy cattle management expert systems can be classified into three categories; advisory expert systems, strategic planning expert systems, and diagnostic expert systems (Spahr *et al.*, 1988). Advisory ES attempt to model the decision making process of an expert and advise the dairy farmer in managerial decisions, for example, sire selection, or culling. Strategic planning ES are used to assist production by proposing strategic management decisions, for example, the prediction of a cow's future performance. Finally, diagnostic ES are used to diagnose equipment malfunctions, or subnormal animal or herd performance. This type of ES can be applied to the interpretation of sensor data (eg. milking machines). The system can indicate when there are problems and suggest recommendations to rectify the situation.

There are many advantages to using expert systems. Expert systems allow the user to trace the reasoning process of a consultation (Domecq *et al.*, 1991). ES can be developed even when information is incomplete, uncertain, subjective, inconsistent, or subject to change (Spahr *et al.*, 1988). In ES the knowledge base and the inference process are kept separate (Nevo *et al.*, 1994) which facilitates the adjustment of the knowledge base as updated knowledge becomes available. Another advantage of ES is that, as opposed to conventional programming, the ordering of the rules is unimportant (McKinion and Lemmon, 1985).

Although there is a great deal of support for expert systems, this does not come without criticism. The development process of an ES requires many years before the ES evolves into a mature product. The duration of this process has been estimated as 5 to 10 years (Smith, 1989, Spahr *et al.*, 1988). The complexity of the subject domain is another consideration. This certainly applies to the dairy sector, since dairy operations are very complex (Doluschitz, 1990). ES for dairy cattle management require knowledge and information from many different sources and this can hinder development. Another critique of ES is the cost of development. McClure (1992) summarized the costs involved in the development of an expert system for the management of apple orchards. The total project cost was $303,160 U.S. over four years. More than a third of this money ($124,160 U.S.) went towards funding of the 5 experts involved in the project who spent a total of 3104 hours contributing their knowledge. Cox (1996) expressed concern that these types of projects siphon-off scarce research money which should be allocated to more research-oriented activities. However, it is conceded that these projects do aid in the establishment of a rapport between the end-users (i.e. the producers) and the researchers, which in turn increases the effectiveness of research and development.

Fuzzy Logic

As was previously mentioned, fuzzy logic is a tool which could benefit expert systems. Fuzzy logic is based on fuzzy set theory which provides a very strict mathematical framework for the manipulation of vague and imprecise information. Zimmermann (1991) specifies three main reasons why fuzzy set theory should be used in expert systems.

(1) Since expert systems are developed by humans for humans, it is most appropriate to create such a system using the "natural" language of humans.

(2) The knowledge base of an expert system is designed from the knowledge of human experts. Since human knowledge is often imprecise or uncertain in nature, it is most appropriate to maintain this knowledge as fuzzy rather than converting to crisp variables and risking a loss of information.

(3) If the knowledge is maintained as uncertain or fuzzy it becomes necessary to find a way to manage this uncertainty. Fuzzy logic is one such tool.

With fuzzy mathematics, sets are used to represent concepts, and membership values are used to denote to what degree a value can be associated to a particular fuzzy set. Fuzzy sets correspond to linguistic labels (sometimes called "qualifiers"). Mathematically, each fuzzy set is denoted by a pair of attributes; the first part being the element, and the second part being the degree of membership to the fuzzy set. Degrees of membership usually range from 0, denoting no membership to a fuzzy set, to 1, representing complete membership. Fuzzy sets are defined by membership functions which give the sets a characteristic shape. The triangular function is usually sufficient to define a fuzzy set (Lee, 1990), but it is the knowledge engineer who usually decides which membership functions to use. An example of a series of fuzzy sets used to describe 305-day milk production can be found in Figure 2.4. From this figure it can be noted that there are no distinct boundaries between sets; the change-over from membership to nonmembership is gradual. Another feature is that as fuzzy sets overlap, the sum of their membership functions can be greater than unity (Grinspan *et al.*, 1994). This is in contrast to classical set theory where an element belongs completely or not at all to a set. The process of qualifying a quantitative value into its fuzzy equivalent is called *fuzzification*. A detailed description of fuzzification and the mathematics of fuzzy logic can be found in Chapter 2 (page 30).

Fuzzy logic has been applied in a number of applications, the most famous being the development of the controller for the subway system in Sendai, Japan (Viot, 1993). Fuzzy set theory has also been applied in such diverse areas as in turbomachinery diagnosis (Siu *et al.*, 1996), the evaluation of cabbage seedling quality (Chen *et al.*, 1994), in conjunction with neural networks to analyze sensory responses for the evaluation of beef steak color (Gao *et al.*, 1994), and in the development of an expert system for determining the optimum time for the transfer of dairy cows from a high feeding group to a lower feeding group (Grinspan *et al.*, 1994).

Knowledge Acquisition

Knowledge acquisition refers the process of acquiring knowledge from an expert source. Research related to the development of procedures for knowledge acquisition originated in the early 80's. The knowledge acquisition procedure was developed as a methodology for the development of expert and knowledge-based systems. Early research in knowledge acquisition from human experts began when traditional software engineering techniques failed to adequately emulate human expertise (Gaines, 1993). Knowledge acquisition can be defined as a procedure for the collection and analysis of information obtained from various sources, including human experts, literature, etc., to form a knowledge-base. A subset of knowledge acquisition is knowledge elicitation. Knowledge elicitation is concerned with obtaining information directly from domain experts (Greenwell, 1988).

Many articles are available which describe knowledge acquisition procedures for expert system development. Many focus on the interview process as the major method for the elicitation of human expertise (Oltjen *et al.*, 1990; Spangler *et al.*, 1989; Greenwell, 1988; Jones *et al.*, 1987; Wright and Ayton, 1987). Other procedures have been developed including: protocol analysis, walkthroughs, questionnaires, expert reports (MDBS, 1991), repertory grids (Greenwell, 1988; Ascough, 1990), twenty questions and card sorting (Jones, 1989).

Many considerations must be addressed before the knowledge elicitation sessions begin. The knowledge engineer should acquire a basic understanding of the domain to be explored by familiarizing him/herself with the subject vocabulary and some of the basic concepts. This is called pre-training (Jones, 1989; Spangler *et al.*, 1989; Jones *et al.*, 1987). Different pre-training methods can be employed, including the consultation of literature, and/or tutorial sessions provided by the domain expert to the knowledge engineer. The latter method provides the advantage of allowing the expert and the knowledge engineer opportunity to get to know one another (Spangler *et al.*, 1989; Jones, 1989). Greenwell (1988) adds that this would help the knowledge engineer obtain the domain vocabulary. However, he expresses the apprehension that pre-training makes the knowledge engineer appear to be more knowledgeable about the subject domain than he/she actually is, which can influence the way in which the expert shares information.

Before interviews are conducted the expert and the knowledge engineer should ensure that they have the same operational goals to avoid loss of development time. This is accomplished by defining how much expertise should be developed, who the final product user will be, and how the product will be presented (Jones, 1989; Greenwell, 1988). Greenwell (1988) states that of the three key players in the development of expert systems - the knowledge engineer, the expert, and the end user - the latter is the most important, and should always be considered throughout the development process.

Because expert system development relies on the quality of knowledge obtained from the expert, care must be taken to ensure the smooth running of the knowledge elicitation sessions. Many researchers have provided suggestions for the interview process:

- The knowledge engineer must show professionalism or risk losing the confidence of the expert (Greenwell, 1988).
- Interviews should be kept to a maximum of two hours to avoid fatigue (Greenwell, 1988).
- Interviews should be conducted away from the expert's workplace to avoid interruptions (Spangler *et al.*, 1989).
- Interviews should be recorded, either by audio tape or videotape, to ensure that everything spoken during the interview is recorded. Video is a better medium if there is a lot of information expressed through visual gestures (Spangler *et al.*, 1989).
- Interviews can be very tiring. The knowledge engineer should be aware of the fatigue threshold and either change the subject area of the interview when the expert appears to tire (to renew the expert's interest) or close the session (Spangler *et al.*, 1989).
- One final consideration is language. Because language is a subjective form of communication, care must be taken to ensure that the expert avoid using false or misused wording. The knowledge engineer must also ensure that the

information is correctly interpreted for use by the system developed (Spangler *et al.*, 1989).

Recent research has shifted the goal of knowledge acquisition from the 'mining' of an experts mind towards a complete modelling of the domain (Allemang and Rothenfluh, 1992; Compton *et al.*, 1993). Some researchers have found that human experts do not provide expertise which is absolute, but rather an interpretation of the world around them. As such, the expert's model is only one of many models which are incorporated into expert systems. Other models can come from sources such as literature, observed skills, observed systems, etc. A more recently evolving area of knowledge acquisition research is the applicability of using computer-based tools for the acquisition of knowledge (Gaines, 1993).

Knowledge Implementation

Knowledge implementation, which is also known as knowledge representation, is the process of developing a knowledge-base from the acquired knowledge. Schneider *et al.*, (1996), describes three methods used to represent knowledge. These are: semantic nets, frames, and production rules. Semantic nets are used to represent propositions. They consist of nodes and links. Nodes are used to designate objects, concepts, or events and the links connect the nodes and describe the relations between nodes. Links are represented with terms such as *is_a*, *a_kind_of*, etc. Inheritance is a feature of semantic nets and therefore, since links are hierarchical, some of the knowledge does not have to be described explicitly. For example, if we consider the following two statements,

the farmer has cows

cows produce milk

a semantic net would make the connection that the farmer's cows produce milk. In this example, *has_a*, and *produce,* which is a property, are both links, while farmer, cows, and milk are all nodes. There are some drawbacks to semantic nets. Semantic nets tend to become very large due to the duplication of nodes. They are also difficult to modify, validation is difficult, and their performance is slow during ES consultation.

The use of frames is another procedure for representing knowledge. A frame is an environment combining declarative and procedural knowledge, and their internal relations. During a consultative process, the expert system selects the most appropriate frame and the frame tries to match itself to the system data. If the frame is not appropriate another frame is selected. Relations are established between frames with the most common being the *parent-child relation*, where the parent represents a class of items and the child represents the subclasses. This is referred to as inheritance. Drawbacks of frames include, large memory requirements, slow performance (although faster than semantic nets), and the need for a very sophisticated inference engine.

As has been previously stated, production rules are the most commonly employed method for representing knowledge. Production rules are often established using a tree format. Tree representation has several advantages. Since it is possible to follow the *path* of the tree, the relations between the rules are easy to see, the inference process is very fast, and it facilitates the explanation of the reasons why a decision was made; it simply becomes necessary to trace the decision-making path backwards. One disadvantage is that it can be difficult to validate the knowledge base, in particular, to ensure that there are no circular rules which may result in unreachable conclusions, or infinite loops. However, if the rule set is relatively small it should not be a problem to manually search through all the individual rules (Harrison, 1991). For larger rule sets there exist other procedures for developing production rules. These include bit matrices, and relational lists (Schneider *et al.,* 1996).

Other methods also exist for the representation of knowledge including: scripts (which are mostly used in natural language systems), logic, and processes (McKinion and Lemmon, 1985). Typically, if knowledge can be represented by at least one of the methods then it should be possible to represent that same knowledge using any of the other methods

(McKinion and Lemmon, 1985). The selection of the most appropriate method depends upon the judgement of the knowledge engineer in determining which method is most efficient in data retrieval and in knowledge deduction.

System Evaluation

In order to ensure that the goal of the DSS has been reached, evaluation must be performed. ES evaluation is performed to measure the system's accuracy and usefulness (Batchelor and McClendon, 1989). Evaluation include the processes of verification and validation (Harrison, 1991). Verification is conducted to ensure the ES performs as intended by searching for missing rules, dead-end IF conditions, mistyped rules, etc. Validation is used to examine whether the structure itself is appropriate and sensitivity analysis is performed to determine the effect that changes to the rules have on the system output (Harrison, 1991, Plant and Stone, 1991). Harrison (1991) describes seven procedures for the validation of a system: face validity, comparison against high-quality performance, minimum standard of performance, the Turing test, post-mortem analysis, field tests, and scope validity. Face validity is performed by a domain expert to simply determine whether the output is reasonable. Since no comparisons are performed, this procedure cannot identify any system weaknesses. Comparison against high-quality performance is performed by comparing the system outputs against a high-quality level of expertise. This expertise can come from a consensus obtained from a panel of experts. Minimum standard of performance is a technique whereby the rate of success of an ES output is measured. For example, an ES which produces correct results 50% of the time might be considered adequate if the human experts themselves also have a 50% success rate. The Turing test, which is also called the 'blind' test, is performed by asking an expert to differentiate between the unlabeled outputs from the ES and from real situations. If the expert cannot distinguish between the two, the ES is considered capable of mimicking the reasoning process of experts. Post-mortem analysis is conducted to determine why an ES has made a wrong recommendation. These failures are examined to determine the areas of weakness and the limitations of the system. Field tests are conducted by the potential users of the system. This places the burden of testing on the users and helps establish the acceptable performance ranges. This procedure also allows for the immediate revelation of any weaknesses in the flow of communication between the system and the user. Finally, scope validity refers to the examination of whether the system encourages the user to examine other possibilities that the system considered but the user overlooked.

There are many considerations which need to be addressed when analyzing the results obtained through validation. Since some ES incorporate the knowledge of multiple experts, such a system might produce recommendations which are better than those ES produced from the knowledge of a single expert (Batchelor and McClendon, 1989). Another consideration is that a comparison between the results from an ES with those obtained from human experts can only be performed if we are absolutely certain that the experts always represent the correct management choice (Willett and Andrews, 1996). Some ES are developed using information that the human expert does not have access to or does not have the time to consider; in these instances the ES can produce results superior to those of a human expert (Harrison, 1991).

III. References

Ascough II, J.C., L.J. Lane, and L. Deer-Ascough. 1990. Interactive knowledge acquisition and expertise transfer: overview and methodologies. American Society of Agricultural Engineers. Paper 90-755.

Allemang, D., and T.E. Rothenfluh. 1992. Acquiring knowledge of knowledge acquisition: A self study of generic tasks. In Wetter, T., Althoff, K.-D., Boose, J., Gaines, B.R., Linster, M., and Schmalhofer, F. (Eds.). *Current Developments in Knowledge Acquisition-European Knowledge Acquisition Workshop 92*. p. 353-372.

Allore, H.G., L.R. Jones, W.G. Merrill, and P.A. Oltenacu. 1995. A decision support system for evaluating mastitis information. Journal of Dairy Science. 78:1382-1398.

Batchelor, W.D. and R.W. McClendon. 1989. Two approaches to knowledge engineering in an insect pest management expert system. American Society of Agricultural Engineers. Paper 89-7081.

Ben-Ari, Y., I. Amir, and S. Sharar. 1983. Operational replacement decision model for dairy herds. Journal of Dairy Science. 66:1747-1759.

Boyd, D.W. and M.K. Sun. 1994. Prototyping an expert system for diagnosis of potato disease. Computers and Electronics in Agriculture. 10:259-267.

Chen, S., R.E. Garrett, N.C. Chen, W.H. Chang, and T.T. Lin. 1994. Evaluation of cabbage seedling quality by fuzzy logic. American Society of Agricultural Engineers. Paper 94-3028.

Compton, P., B. Kang, P. Preston, and M. Mulholland. 1993. Knowledge acquisition without analysis. In Aussenac, N., Boy, G., Gaines, B, Linster, J.-G., and Kodratoff, Y. (Eds.) *Knowledge Acquisition for Knowledge-Based Systems-Acquisition-European Knowledge Acquisition Workshop 93*, p. 277-299.

Cox, P.G. 1996. Some issues in the design of agricultural decision support systems. Agricultural Systems. 52:355-381.

Dentine, M.R., B.T. McDaniel, and H.D. Norman. 1987. Comparison of culling rates, reasons for disposal, and yields for registered and grade Holstein cattle. Journal of Dairy Science. 70:2616-2622.

Doluschitz, R. 1990. Expert systems for management in dairy operations. Computers and Electronics in Agriculture. 5:17-30.

Domecq, J.J., R.L. Nebel, M.L. McGilliard, and A.T. Pasquino. 1991. Expert system for evaluation of reproductive performance and management. Journal of Dairy Science. 74:3446-3453.

Gaines, B.R. 1993. Modelling and extending expertise. In Aussenac, N., Boy, G., Gaines, B, Linster, J.-G., and Kodratoff, Y. (Eds.) *Knowledge Acquisition for Knowledge-Based Systems-Acquisition-European Knowledge Acquisition Workshop 93.* p. 1-22.

Gao, X., J. Tan, and D. Gerrard. 1994. Applying fuzzy logic and neural network in analysis of sensory responses. American Society of Agricultural Engineers. Paper 94-3558.

Greenwell, M. 1988. Knowledge Engineering for Expert Systems. Ellis Horwood Ltd. Chichester, UK.

Greer, J.E., S. Falk, K.J. Greer, and M. J. Bentham.1994. Explaining and justifying recommendations in an agricultural decision support system. Computers and Electronics in Agriculture. 11:195-214.

Grinspan, P., Y. Edan, H.E. Kahn, and E. Maltz. 1994. A fuzzy logic system for dairy cow transfer between feeding groups. Transactions of the ASAE. 37:1647-1654.

Harrison, S.R. 1991. Validation of agricultural expert systems. Agricultural Systems. 35:265-285.

Jain, D.K., U.S. Tim, and R.W. Jolly. 1995. A spatial decision support system for livestock production planning and environmental management. Applied Engineering in Agriculture. 11:711-719.

Jones, J.W., P. Jones, and P.A. Everett. 1987. Combining expert systems and agricultural models: A case study. Transactions of the ASAE. p. 1308-1314.

Jones, P.H. 1989. Knowledge acquisition. In Barrett, J.R., and D.D. Jones. (Eds.). Knowledge engineering in agriculture. American Society of Agricultural Engineers. St. Joseph, MI. p. 33-46.

Kennedy, J.O.S. and A.W. Stott. 1993. An adaptive decision-making aid for dairy cow replacement. Agricultural Systems. 42:25-39.

Laacke, R.J. 1995. Building a decision support system for ecosystem management: KLEMS experience. AI Applications. 9:115-127.

Lacroix, R., K. Wade, R. Kok, and F. Hayes. 1994. Functional characteristics of a decision support system for dairy cattle management. American Society of Agricultural Engineers. Paper 94-3027.

Lee, C.C. 1990. Fuzzy logic in control systems: fuzzy logic controller (Part 1). IEEE Transactions on Systems, Man, and Cybernetics. 20: 404-418.

Levins, R.A., and M.A. Varner. 1987. An expert diagnostic aid for reproductive problems in dairy cattle. Computers and Electronics in Agriculture. 2:47-56.

Lokhorst, C. and E.J.J. Lamaker. 1996. An expert system for monitoring the daily production process in aviary systems for laying hens. Computers and Electronics in Agriculture. 15:215-231.

Mainland, D.D. 1994. A decision support system for dairy farmers and advisors. Agricultural Systems. 45:217-231.

McClure, J.E., M. Tomasko, and C.H. Collison. 1993. BEEAWARE, an expert system for honey bee diseases, parasites, pests, and predators. Computers and Electronics in Agriculture. 9:111-122.

McClure, J.E. 1992. Cost of developing an expert system in agriculture. AI Applications. 6:67-73.

McDaniel, B.T. 1994.Genetic evaluations of secondary traits in dairy cattle. Proceedings of the 5th World Congress on Genetics Applied to Livestock Production. 17:59-60.

McKinion, J.M. and H.E. Lemmon. 1985. Expert systems for agriculture. Computers and Electronics in Agriculture. 1:31-40.

MDBS. 1991. Building expert systems using GURU. Micro Data Base Systems, Inc. Lafayette, Ind.

Monardes, H. 1992. La réforme des vaches. Le Producteur de Lait Québécois. Mars. 1992. p. 14-15.

Nevo, A., R Oad, and T.H. Podmore. 1994. An integrated expert system for optimal crop planning. Agricultural Systems. 45:73-92.

Oltjen, J.W., Selk, G.E., Burditt, L.G., and Plant, R.E. 1990. Integrated expert system for culling management of beef cows. Computers and Electronics in Agriculture. 4:333-341.

Pellerin, D., R. Levallois, G. St-Laurent, and J.-P. Perrier. 1994. LAIT-XPERT VACHES: An expert system for dairy herd management. Journal of Dairy Science. 77:2308-2317.

Pellerin, D. 1993. L'Elevage des genisses: La moderation a bien meilleur cout. Symposium sur les Bovins Laitiers. 28 Octobre 1993. Conseil des Productions Animales du Québec. p. 13-57.

Plant, R.E. and N.D. Stone. 1991. Knowledge-Based Systems in Agriculture. McGraw-Hill, U.S.A. 365 pp.

Rauscher, H.M. 1995. Natural resource decision support: theory and practice. AI Applications. 9:1-2.

Rogers, G.W., J.A.M. Van Arendonk, and B.T. McDaniel. 1988. Influence of involuntary culling on optimum culling rates and annualized net revenue. Journal of Dairy Science. 71:3463-3469.

Schneider, M, A. Kandel, G. Langholtz, and G. Chew. 1996. Fuzzy Expert System Tools. John Wiley and Sons. England. 198 pp.

Secher, B.J.M., L.N. Jorgensen, N.S. Murali, and P.S. Boll. 1995. Field validation of a decision support system for the control of pests and diseases in cereals in Denmark. Pesticide Science. 45:195-199.

Siu, C.C., Q. Shen, and R. Milne. 1996. A fuzzy expert system for turbomachinery diagnosis. First Online Workshop in Soft Computing. Aug. 19-30, 1996. http://www.bioele.nuee.nagoya-u.ac.jp/wsc1

Smith, T.R. 1989. The potential application of expert systems in dairy extension education. Journal of Dairy Science. 72:2760-2766.

Spahr, S.L., L.R. Jones, and D.E. Dill. 1988. Expert Systems-their use in dairy herd management. Journal of Dairy Science. 71:879-885.

Spangler, A.M., C.D. Ray, and K. Hamaker. 1989. Knowledge acquisition for expert system development. Computers and Electronics in Agriculture. 4:23-32.

Stewart, H.M., E.B. Burnside, and W.C. Pfeiffer. 1978. Optimal culling strategies for dairy cows of different breeds. Journal of Dairy Science. 61:1605-1615.

Stewart, H.M., E.B. Burnside, J.W. Wilton, and W.C. Pfeiffer. 1977. A dynamic programming approach to culling decisions in commercial dairy herds. Journal of Dairy Science. 60:602-617.

Thieme, R.H., J.W. Uhrig, R.M. Peart, A.D. Whittaker, and J.R. Barrett. 1987. Expert system techniques applied to grain marketing analysis. Computers and Electronics in Agriculture. 1:299-308.

Van Arendonk, J.A.M. 1988. Management guides for insemination and replacement decisions. Journal of Dairy Science. 71:1050-1057.

Viot, G. 1993. Fuzzy logic: concepts to constructs. AI Expert. Nov. 1993. p. 26-33.

Willett, M.J. and P.K. Andrews. 1996. Knowledge-based systems for fruit and vegetable production management. AI Applications. 10:75-92.

Wright, G. and P. Ayton. 1987. Eliciting and modelling expert knowledge. Decision Support Systems. 3:13-26.

Young, G.B., G.J. Lee, D. Waddington, D.I. Sales, J.S. Bradley, and R.L. Spooner. 1983. Culling and wastage in dairy cows in East Anglia. Veterinary Record. 113:107-111.

Zimmermann, H.J. 1991. Fuzzy Set Theory - and Its Applications (2nd edition). Kluwer Academic Publishers. Norwell, Massachusetts, USA. 399 pp.

IV. RESEARCH OBJECTIVES

The objective of this project was to develop a decision-support system to recommend dairy cattle culling decisions. It was envisaged that this study would be conducted in three stages. The goal of the first stage was to investigate the use of a fuzzy logic approach in the development of DSS for the area of dairy-cattle breeding. This investigation was done through the development of a prototype decision-support system that would make culling recommendations for individual cows, based on test-day records. The specific objectives were to: 1) develop the prototype DSS, 2) analyse its performance as well as other factors affecting its performance, and 3) establish a basis for the development of more complete DSS for breeding decisions.

Because it was felt that a methodological framework was required to pursue the development of DSSs for dairy cattle breeding, and because a framework suiting our needs was not found, the decision was made to create one. The second stage of this project was, therefore, to establish a more thorough methodology for the development of fuzzy logic based decision-support systems. The development of this framework was accomplished through an analysis of the prototype previously developed, in combination with existing literature. The framework encompasses the groundwork necessary at project outset, the knowledge elicitation and implementation processes, the defuzzification procedure, and the resolution of "conflict" when multiple-experts are involved.

Based upon the results of the prototype DSS developed in the first stage, it was determined that the creation of a second-generation prototype should be pursued. The objective of this study was to develop a second generation culling decision-support system prototype. Specifically, the objectives were:

- to acquire expert knowledge using the procedures outlined in the framework developed in the second stage, expanding the expertise to include the knowledge of six experts from three different domain areas,

- to develop individual DSS modules for each expert, whereby each expert's knowledge would be maintained independently from the knowledge of other experts,
- to examine the possibility of integrating the six modules into a single aggregated DSS,
- and to pursue the evaluation of the applicability of fuzzy logic in DSS.

CHAPTER 1

Fuzzy Logic-Based Decision-Support Systems for Culling of Cows

R. LACROIX, M. STRASSER, R. KOK and K.M. WADE

ABSTRACT

In order to investigate the use of fuzzy logic in decision-support systems (DSSs) for dairy cattle breeding, a prototype software system was developed. The objectives were to determine the advantages and limitations of fuzzy logic for this type of application, and to establish a basis for the development of more complete DSSs in the future. The goal of the prototype DSS was to make culling decisions on the basis of monthly production data. During the development phase, three experts in the area of animal breeding were interviewed. The final version comprised three rule sets which considered a total of five input variables. The membership functions for most of the input variables were made herd-specific. Results showed that the use of fuzzy sets could increase the flexibility and adaptivity of rule-based expert systems. The same rule sets were appropriate under various scenarios (e.g., herds, regions and breeds) with inferences being made specific to each context, by adjusting the membership functions associated with the fuzzy sets. Results also showed that the inferences from fuzzy sets could be used as an alternative to methods currently used for within-herd cow rankings. Fuzzy sets appear to facilitate the development of expert systems and might require a smaller number of rules than traditional approaches which mimic breeders' reasoning processes.

(**Key words:** dairy-cattle, culling, decision-support systems, fuzzy logic)

Abbreviation key: DSS = decision-support system, **DHIA** = Dairy Herd Improvement Agency,

MV = membership value, **RS** = rule set.

INTRODUCTION

The animal breeding programme is an integral part of any efficient dairy enterprise, the major components of which involve decisions concerning culling, replacement and mating. Culling decisions are concerned with achieving a balance between genetic progress for specific traits of economic importance and an animal's length of productive life, while replacement and mating decisions cover such issues as age structure of the herd and the identification of traits which need to be improved in the overall genetic profile of animals in the herd. On an individual cow basis, breeding involves decisions about, for example, optimal culling time, if and how she should be replaced and which sire should be mated with her to better improve the genetic makeup of the next generation. In general, breeding decisions are complex and must take into account a large number of interrelated factors. Therefore, they require the consideration of many information sources, part of which is collected on the farm (e.g., milk production, fertility and conformation), the other portion being made available through external agencies such as Artificial Insemination units, Dairy Herd Improvement Agencies (DHIAs) and Breed Associations (e.g., genetic proofs for conformation and pedigree information). Potentially, during a decision-making process, the larger the number of factors, the better the conclusions can be.

While optimal breeding decisions require the processing of large volumes of information, the human ability to carry out this task is limited. Humans are able to make complex decisions through information analysis and reasoning, but, within a certain period of time, they can analyze only a finite number of data contained in tables and figures. Also, at some point during any analysis, humans become saturated and can no longer absorb further information. In contrast, computers are particularly good at rapidly and endlessly carrying out well-structured, procedural tasks; they are good information crunchers. Therefore, humans and computers both possess different strengths, and, when trying to make an optimal decision, the best approach would seem to consist of combining computerized treatment and human thinking. With this approach, the computer software components form decision-support systems (DSSs), which act as information pre-digesters, and which establish the basis for final decisions by human managers. This approach can be applied to dairy cattle breeding, whereby the role of DSSs would be, for example, to identify potentially problematic areas, to make preliminary diagnostics and to formulate recommendations. The DSSs would then carry out lower-level and repetitive analysis tasks, and would leave higher level/final decisions to the farm manager.

In order to help the final decision-making process of humans, it is important to develop as much as possible the analytical ability of DSSs. A good approach for this consists of capturing the expertise of specialists in well-defined, narrow areas, and then embodying it in software modules for addition to DSSs. These modules, traditionally called "expert systems", can become important components of DSSs for dairy cattle breeding, and can be particularly helpful in diagnosing problems and suggesting recommendations for solving them. However, in order for expert systems to be able to mimic the reasoning and decision-making processes of specialists, they must be able to deal with vagueness, ambiguity and uncertainty (4, 12). Many mathematical tools exist to help developing such software, e.g., tools based on confirmation, Bayesian probability and fuzzy set theories (3, 5, 7, 13). In the last decade, much attention has been given to fuzzy logic in various economic sectors, and many commercial applications have been developed, based on this approach (11). An advantage of fuzzy logic is that it allows for approximate reasoning and decision-making based on vaguely defined, linguistic variables, which generally characterize the decision-making processes of experts. Since fuzzy logic has been applied successfully in many agricultural areas (1, 2, 4, 10), it seems reasonable to consider its use in the area of dairy cattle management.

The goal of this research was to investigate the use of a fuzzy logic approach in the development of DSSs for the area of dairy-cattle breeding. This investigation was done through the development of a prototype decision-support system that would make culling recommendations for individual cows, based on test-day records. The specific objectives were to: 1) develop the prototype DSS, 2) analyze its performance as well as other factors affecting its performance, and 3) establish a basis for the development of more complete DSSs for breeding decisions.

MATERIALS AND METHODS

Fuzzy Sets

In decision-making processes, experts often use qualitative terms in describing something about which they are reasoning. For example, an expert may qualify the milk production level of a specific cow as being 'medium' and her fertility as being 'low'. An important aspect is that no sharp boundary exists between the qualifiers used. For example, if an expert assesses a cow's average calving interval of 392 days as 'medium', he will most likely not describe a calving interval of 393 days as 'long'. At the same time, many experts might agree that an average calving interval of 420 is *definitively* 'long' and a calving interval of 390 days is definitively 'medium'. Fuzzy mathematics can be used to deal with such situations in a quantitative manner.

Fuzzy mathematics is based on fuzzy sets, which correspond to the qualifiers employed by specialists. Each possible qualifier used to describe a situation or an entity corresponds to one fuzzy set, with a series of fuzzy sets used to cover all possible qualification levels (e.g., from 'very short' to 'very long' in the case of average calving interval). In contrast to classical set theory, where an element belongs either completely or not at all to a specific set (e.g., the set of cows with an average calving interval larger than 400 days), fuzzy set theory allows elements to belong ***partially*** to different sets. For this reason, once fuzzified, a numerical value is characterized by one or more fuzzy sets (i.e., the sets to which they belong), and a degree of membership in each of these sets.

An example of fuzzy sets which describe 305-day milk yield is shown in Figure 1.1. Here, all cows with a milk production less than 6,000 kg belong entirely to the set 'VeryLow'. Between 6,000 and 7,000 kg, the cows belong to two sets: 'VeryLow' and 'Low'. As production increases from 6,000 kg, the degree of membership in the set 'VeryLow' decreases, while the membership in the set 'Low' increases. At 7,000 kg, the production is no longer 'VeryLow' and is definitively 'Low'. In this example, five sets are used, but more or less could be used. The number of sets depends on the desired precision and/or on the degree of fuzziness. In this example, the shapes of the sets are triangles and trapezoids; these shapes are commonly used because they ease numerical computation. However, other shapes could also be used. Also, in this example, any production level can be mapped to a maximum of two sets, and the degrees of membership always sum to one; in theory, an element could be mapped to more than two sets, and the degrees of membership could sum to a value which differs from one. Generally, the number of sets, their shape and the numerical values that position them along the X-axis should be chosen so as to reflect accurately the experts' knowledge.

During a specific inference, the mapping of a numerical, crisp value to its fuzzy equivalent (i.e., the determination of the fuzzy sets to which it belongs and the associated degrees of membership) is called the "fuzzification" process. For example, in Figure 1.1, a milk production of 7,338 kg is classified simultaneously as 'Low' with a membership value (MV) of 0.66, and as 'Average' with a degree of membership of 0.34. This is denoted as:

MilkProd = (Low, 0.66), (Average, 0.34)

where "MilkProd" is the 305-day milk yield of a specific dairy cow.

Fuzzy Logic

Fuzzy mathematics makes it possible to deal numerically with fuzzy sets, mainly through operations performed on their degree of membership. In the past three decades, research has been carried out to apply fuzzy sets to various branches of traditional mathematics. Fuzzy mathematics has now become an extensive field of study, and applications vary, for example, from fuzzy clustering to fuzzy linear programming (13). An important application of fuzzy set theory for decision-making and approximate reasoning is fuzzy logic, which is concerned with drawing inferences using rules constructed with fuzzy variables.

In fuzzy logic, an inference process consists of determining the MV of the fuzzy variables contained in a rule's conclusion. This value is a function of the degree of truth of the premise. The truth of the premise is a function of i) the degree of membership associated with the values of the fuzzy variables contained in the premise and ii) the logical operator(s) that link(s) these variables. The two basic operators are AND and OR, which respectively correspond to the intersection and union operations of set theory. Their use in fuzzy inferencing is illustrated below. Consider the set of hypothetical rules in Figure 1.2, with the production of the cow as previously discussed in the fuzzification example (Figure 1.1) and with a reproductive efficiency classified as VeryLow at 0.39 and Low at 0.74:

MilkProd = (Low, 0.66), (Average, 0.34)

ReprodEff = (VeryLow, 0.39), (Low, 0.74)

where "ReprodEff" is the reproductive efficiency.

The premise of RULE 1 is composed of two conditions that are linked with an **OR** operator. Using extension principles, the degree of truth of such a statement can be theoretically established using various methods (13). However, the most common method consists of taking the *maximum* MV of the fuzzy sets involved. In this case, since one of the milk production values is (Low, 0.66) and one of the reproductive efficiency values is (Low, 0.74), the premise of RULE 1 is true to a degree of 0.74. Consequently, the degree of membership of the variable "Culling" to the fuzzy set 'yes' is:

Culling = (yes, 0.74)

In RULE 2 and RULE 3, an **AND** operator links two conditions. Again, using extension principles, the truth of such a statement might be determined by using various methods, but the most common procedure associated with the **AND** operator consists of taking the *minimum* MV of the fuzzy sets in the statement. When this operation is applied to RULE 2, it is now established that:

Culling = (yes, 0.34)

By applying the same operation in RULE 3, it is concluded that:

Culling = (no, 0.34)

During an inference process, many rules in a rule set can be fired since fuzzy variables can possess more than one value; in this example, three rules were fired. For this reason, different rules may assign a variable to the same fuzzy set but with different MVs. For example, RULE 1 and RULE 2 led to the conclusion that the variable "Culling" belonged to the fuzzy set 'yes' with MVs of 0.74 and 0.34 respectively. In this case, it is necessary to assign a final degree of membership to the set, and this can be done with different methods. The simplest method consists of assigning the maximum degree of membership (i.e., the common fuzzy union operation). Other methods, derived from extension principles, or coming from confirmation theory, may also be used. For example, it may sometimes be convenient to use methods that are based on the supposition that two conclusions in agreement confirm and reinforce each other. One such method, called the 'probability sum method' was proposed for confirmative certainty algebra in expert systems (6). Using this method in the previous example, the new degree of membership to the set 'yes' for the variable "Culling" would be:

$$\mu_{\text{culling, yes}} = 0.74 + 0.34 - (0.74 \cdot 0.34) = 0.83$$

where $\mu_{\text{culling, yes}}$ is the degree of membership to the set 'yes'.

Thus, with this method, the firing of Rules 1 through 3 would lead to the following values for "Culling":

Culling = (yes, 0.83), (no, 0.34).

At the end of a fuzzy inference process, the output variables generally possess more than one value, as in the previous example. Each value is composed of a fuzzy set and a numerical value representing the degree of membership to this set. However, a single and crisp value is usually needed in order to generate an overall conclusion. The process by which this single, crisp value is obtained is called "defuzzification". Various defuzzification techniques exist, depending on whether the output variable is continuous or discrete. When dealing with continuous variables, a single numerical value is usually needed at the end of the inference. A common method to obtain such a value is called "centroid defuzzification" (11). In the case of a discrete variable, the approach can simply consist of choosing the fuzzy set with the largest MV. Using this approach, the final conclusion of the previous example would be a culling decision of 'yes' (i.e., 0.83 > 0.34).

Development of the Software Prototype

In this project, the fuzzy logic-based DSS prototype for culling recommendations was constructed using knowledge engineering techniques based on readings and interviews with three local specialists. The knowledge acquisition from the specialists was done in three stages: 1) informal discussion, 2) variable definition and rule formulation, and 3) determination of membership functions. In the initial stage, interviews were conducted with the experts to discuss the factors (variables) on which the DSS should focus, given available data, when making culling recommendations for individual cows. Subsequent interviews were conducted to determine the links between variables which the knowledge engineer used to design the architecture of the prototype DSS. The experts were also asked which descriptors (i.e., fuzzy sets) they would use for each variable, and which conclusions they would draw from various sets of conditions (i.e., from various combinations of fuzzy sets representing different variables). This led to the development of the rules. When the base system was developed, numerical values for all variables considered by the DSS were presented to the experts. These values represented individual test-day records retrieved from a data set supplied by the local DHIA (Québec Dairy Herd Analysis Service). The experts were asked 1) to qualify the numerical values and 2) to make culling decisions based on these data. From this information it was possible to determine the membership functions and further fine-tune the DSS.

A modular approach was adopted for the implementation of the prototype DSS and the knowledge base was encoded into several sets of rules. The software system was developed with GURU 3.0 (8), running under the operating system OS/2. GURU is an integrated product which includes a rule-based expert system shell, a data base management system, a procedural language and spreadsheet capacity. GURU also allows the manipulation of fuzzy (multi-valued) variables and contains many tools for the manipulation of fuzzy sets and their associated MVs. The expert system shell is conceived for making inferences using fuzzy variables and certainty factors, and various inference methods are available (backward, forward and mixed chaining, rule firing based on cost or priority criteria, etc.).

Performance Analysis

Once the software prototype was developed and implemented, some numerical experiments were carried out to evaluate its performance within various contexts. The main objective was to gain some insight regarding the impact of input values and membership functions on the DSS's outputs. There was no attempt to modify or evaluate the importance or impact of the internal constituents of the DSS (i.e., variables and rules), which were considered as static. The experiments were performed with individual Holstein test-day records representing 30 herds and 804 cows after edits (for example, all cows with incoherent data or without 305-day production values were removed). The first step of the performance analysis consisted of studying the inference processes produced by the DSS. This was done in detail for three arbitrarily chosen herds (Herds 9, 21 and 30) during the system development and the validation period. The second step consisted of analysing the overall results of the inference process for all 30 herds.

RESULTS

Overall Decision-Making Process

Local experts determined that a decision to cull should be based on various factors such as production level, age of cow, conformation, health, etc. However, given available data (and for the purpose of this prototype), it was decided that the decisions of the DSS would be based only on the production level of the cow, her reproductive efficiency and her lactation number. Also, the DSS was to be applicable only to multiparous cows. The variables involved in the overall decision-making and their interrelationships are shown in Figure 1.3. Most variables were defined as fuzzy and the fuzzy descriptors, listed in Table 1.1, were used. Using a modular approach, the architecture of the DSS was organized so that the overall decision-making process would be carried out by three sets of rules. The three rule sets were respectively used to evaluate the reproductive efficiency, to make a culling decision (in fuzzy terms) and to defuzzify the culling decision after analysis.

The reproductive efficiency of the cow was evaluated by the first rule set (RS 1) using calving interval, days to first breeding and number of breedings. For the three input variables, the membership functions used to determine the sets in which a certain cow belonged and the degree of membership to each set were specific to each herd; they were constructed relative to the herd average values for each variable, as described below. Twenty-seven rules were developed to cover all possible combinations of fuzzy sets associated with the three input variables. These rules are partially listed in Appendix 1.1. Using these rules, RS 1 characterized reproductive efficiency with one of its four descriptors. The degree of membership in each set was derived during an inference using the minimum method for **AND** operations (there were no **OR** operations in the rules). The confirmative calculations across the rules were done using the probability sum method.

The second rule set (RS 2) required the values for reproductive efficiency, lactation number and production index to make a culling decision. The lactation number had two possible values: 'less than three' or 'three and greater'. The production index was a linear combination of milk, fat and protein production. This was done since the relative economic importance of each dairy component varies from region to region and from time to time. The production index was calculated using equation 1:

$$\text{ProdIndex} = \alpha * \text{Milk} + \beta * \text{Fat} + \gamma * \text{Protein} \tag{1}$$

where:

ProdIndex:	Production Index	[kg]
Milk:	Average 305-Day milk production	[kg]
Fat:	Average 305-Day fat production	[kg]
Protein:	Average 305-Day protein production	[kg]

α, β, γ: Weighting factors for milk, fat and protein, respectively

For this prototype, the values for α, β and γ were 1, 10 and 20, respectively, which represent current economic values for each of the three production traits. Once fuzzified, the production index was described using five qualifiers and, again, the membership functions associated with the five sets were different for each herd and were determined using the herd's average production values. RS 2 was composed of 10 rules with two possible outputs: 'yes' and 'no'. The 'Culling' variable's degree of membership for each of the two possible values was determined during the inference, again applying the minimum operation and the probability sum method. Rules composing RS 2 are listed in Appendix 1.2. It should be noted that reproductive efficiency (in conjunction with lactation number) was *only* considered in decisions concerning cows with a production index of medium. In previous versions of RS 2, reproductive efficiency was also considered for other categories of production index, but, during the validation process, experts finally concluded that culling should be based mostly on production, and that reproductive efficiency must be used only to discriminate between medium producing cows.

Since inferences with the previous rule sets led, in most cases, to two values ('yes' and 'no'), with various degrees of membership for the two descriptors, a way to defuzzify this conclusion was required for the DSS to furnish a specific, crisp, recommendation. It was decided to use a rule-based approach for this specific situation, and a third rule set (RS 3) was added to the DSS. RS 3 was composed of four rules and produced one of three possible conclusions: 'yes', 'no' and 'unknown'. For example, RS 3 considered the difference between the degree of membership for the two possible values of the variable 'culling'; if the difference was too small, the decision was 'unknown' (see rules in Appendix 1.3).

Membership Functions

During the acquisition phase, the specialists consistently requested herd-average values when trying to make a culling decision for any specific cow. Early on in that phase, it was realized that the exact meaning of the descriptors used by the specialists, when presented with numerical values (e.g., calving interval values), varied from one herd to another. At the same time, each combination of descriptors (in rules' premises) was consistently leading to the same conclusion, independently of the herd average values. The use of fuzzy logic then appeared to be ideal to mimic the reasoning of the specialists in this context. Indeed, it was possible to use the same rule set with different herds, making the inference herd-specific by adjusting the membership functions associated with the various fuzzy sets.

In the present DSS, it was then assumed that, for the fuzzy input variables, 'normal' or 'medium' or 'average' conditions for any specific herd were those corresponding to the average values for that herd. The fuzzy sets and their respective membership functions were then defined relative to these average conditions. This was done for the variables calving interval, days to first breeding and production index. Since the data for the variable 'number of breedings' were not considered reliable (e.g., number of breedings were not always reported), its default fuzzy sets were the same for every herd.

All fuzzy sets were defined using triangles and trapezoids. This approach was considered sufficient for the purpose of the prototype DSS, and was used to simplify numerical computations. The MVs for any set varied between 0 and 1, and the sum of the MVs associated with the fuzzy sets at any point always equalled 1. Except for 'number of breedings', the middle fuzzy sets were centered on the herd average values. The position of the other sets was determined by what is referred to as a 'critical point', which corresponds to the summit for triangles or the point of discontinuity in the upper portion of trapezoids. For 'calving interval' and 'days to first breeding', the critical points were set by default at plus and minus 30 days, and at plus and minus 10 days, respectively. This is illustrated for 'calving interval' in Figure 1.4, where the central value (409 days) corresponds to the average calving interval for a specific herd. In this example, a calving interval of 392 days (for a hypothetical cow belonging to this herd) would be simultaneously considered as (short, 0.57) and (average, 0.43). For the production index, the critical points were set to plus and minus one and two increments of 2000 kg (Figure 1.5). For number of breedings, the position and the shape of the sets were the same for all herds; critical points were at 1, 1.5 and 2 breedings.

Performance Analysis

An example of the inference process for an individual cow is shown in Figure 1.6. The inputs to each of the three RSs are also listed before each partial inference, and the rules are displayed in the order in which they were fired. It should be noted that many rules may be fired during an inference process. Also, some output variables may contain more than two values after an inference (e.g., reproductive efficiency). The effect of the confirmative calculation with the probability sum method can be seen in Figure 1.6. For example, the firing of RULE 17 of the first RS should result in a reproductive efficiency value of (poor, 0.66). However, since reproductive efficiency has already been determined as being (poor, 0.18) after firing RULE 15, the degrees of membership are combined to produce a new value of 0.72.

The inference results obtained for the individual cows of three specific herds (9, 21 and 30) are listed in Tables 1.2 through 1.4. The results are presented in a decreasing order of the MV associated with the 'no' value. To a certain extent, the degrees of membership to 'yes' and 'no' are complementary: low degrees of membership to 'yes' generally correspond to high degrees of membership to 'no', and vice-versa. When considering the final culling decisions of the DSS (last column of Tables 1.2 through 1.4), it can be observed that the proportion of 'yes' cases is similar for Herds 9 and 21 (26% and 24%, respectively). This is true even though average production indices for these two herds differ considerably, as indicated in Table 1.5. However, for Herd 30, the proportion of 'yes' cases is close to 50%, even though the critical values that are used are specific to this herd. Specifically, a 'no' decision was produced for all cows (except one) in this herd where the production index was larger than the average production index for the herd (19223 kg - see Table 1.5), and a 'yes' decision was produced for all cows (except one) with a production index lower than the average production index for the herd. This is different to what occurs in Herds 9 and 21, where eight cows and five cows, respectively, with a production index lower than the herd average were not categorized as 'yes'. The disparity in the proportions of 'yes' cases is due to the difference in the distribution of the individual production indices within the herds, to the membership functions associated with production index, and to the rules which compose RS 2 in the DSS. In the case of Herd 30, the production index of the cows was very variable and the deviations were generally quite large about the herd average. Due to this variability on both sides of the herd average value (which is indicated by a high standard deviation in Table 1.5), the degree of membership to the set 'medium' was generally not very high (i.e., from 0.05 to 0.61). Consequently, even when reproductive efficiency and lactation were considered in the inference process of 14 cows out of 21 (or 67%) for Herd 30, their influence on the final conclusions were small. This is due to the use of the 'minimum' operation in the premise of the rules; when the degree of membership to 'medium' for production index is very small, then the probability that it be the element limiting the degree of truth of the whole premise is high. Therefore, for Herd 30, the discrimination among cows was based mostly on production indices, which were evenly distributed about average herd production indices. In the case of Herds 9 and 21, for cows with a lower production index than the herd average and which were not categorized as 'yes', the decisions were favourably influenced by the reproductive efficiency, which was good to excellent, or by the lactation number when reproductive efficiency was poor or unsatisfactory (i.e., the lactation number was than less or equal to three in this case). For all those cows, the degree of membership to the set 'medium' for production index was generally large (from about 0.57 to 0.97) and, consequently, the weight of reproductive efficiency in the final decision was large.

The inference results for all 30 herds are presented in Table 1.6. The culling rate (i.e., when 'yes' is recommended) varies from 14% to 49%, with an average of 31%. Note that this average value is close to what is often suggested for dairy farms. The proportion of 'no' fluctuates between 44 and 75%, with an average value of 62%. The percentage of 'unknown' cases varies between 0 and 24, the average being 8%. Since it was previously observed that the distribution of final decisions (e.g., the overall culling rate) was influenced by the distribution of the input values about herd averages, it was hypothesized that a relationship may exist between the distribution of the outputs and the standard deviation of the input variables. Thus, the proportion of 'no' and 'yes' cases was plotted against the standard deviation of the production index for the 30 herds (Figure 1.7). A certain trend can be observed (higher standard deviations often lead to larger proportions of 'yes' cases and to smaller proportions of 'no' cases), even if the relationships are not very strong (linear correlation coefficients are .54 and .31 for 'yes' and 'no' cases respectively). This can be explained by the fact that, the higher the standard deviation of the production index, the higher the probable number of cows in a herd for which reproductive efficiency is not considered (the correlation coefficient between standard deviation of the production index and percentage of cows for which the deviation of the production index about herd average is larger than 2000 kg is 0.76). Also, production index alone is used to discriminate cows for which the deviation of the production index about herd average is larger than 2000 kg. Since, on average, the cows are fairly well distributed about herd averages, herds with high standard deviations of the production index will tend to generate final results that are more evenly distributed (i.e., the number of 'yes' cases will be closer to the number of 'no' cases). For other variables such as calving interval, similar trends were not detected. This absence of relationship is probably due to the fact that a variable such as calving interval affects the final inference results indirectly through reproductive efficiency; also, reproductive efficiency was only considered in the case of medium producing cows. However, it should be indicated that, on average, 71% of the cows had a production index deviation about herd average of smaller than 2000 kg. Consequently, on average, reproductive efficiency was considered in the inference process for 71% of the cows, and production index alone does not explain the inference results for those cows.

DISCUSSION

Although the prototype DSS was narrow in scope and considered only a relatively few number of variables in making its decisions, its construction allowed for the determination of a series of factors to be considered during the development of a fuzzy logic based DSS. It also allowed for some advantages and limitations of fuzzy logic to be examined. An important point that was observed is that fuzzy logic facilitates the incorporation of knowledge into a software system because it permits the use of linguistic qualifiers such as those employed by the domain experts. It also appears that, with fuzzy sets, the lines of reasoning, used by specialists, can be reproduced remarkably well with a smaller number of rules than with approaches based on crisp sets. For example, three to five qualifiers were sufficient to describe factors such as those used in the DSS. Combining these sets to compose rules and using membership functions to numerically express the meaning of the sets, allowed a system to be developed which could theoretically lead to an infinite number of possible outputs (before defuzzification). Indeed, the use of fuzzy sets with the input variables could lead to any degree of membership to the two sets associated with the output variable. In order to obtain a similar behaviour with crisp sets, a larger number of rules would have been required so as to cover many possible cases within the variation ranges of the input variables. If the number of rules, required to embed an area of expertise, is considerably lower with fuzzy sets than with crisp sets, then the implementation and debugging of fuzzy expert systems could potentially be easier, while leading to lower development time and costs as well. For similar reasons, fuzzy expert systems may be easier to maintain and upgrade than traditional ones.

An interesting observation from this study was that DSSs, based on fuzzy sets, are flexible and can easily be adapted to various contexts without changing the rules contained in the knowledge base. The parameters that characterize specific inference processes (i.e., MVs and functions) are external to rule sets and can be defined at each inference. For example, the same sets of rules could be used for farms with different levels of productivity or objectives, the DSS being adapted to each farm by modifying only the membership functions associated with the fuzzy sets. This means that a fuzzy logic based DSS could easily be adapted to different regions or to various breeds of cattle. This approach also potentially increases the lifetime period of a DSS compared to a situation where crisp variables are incorporated into rules. Such a DSS might arguably remain valid despite changes due to genetic improvement or even better management.

Concerning the analyses performed with the prototype DSS under various conditions, the analyses permitted insight into the intrinsic characteristics of the DSS itself, and the general use of fuzzy systems. Analyses displayed the fact that the results were influenced by many factors. For example, the inference processes relied heavily on membership functions, as well as on the values of the input variables. Rules also played a key role in the inference process since, in conjunction with membership functions, they transformed the inputs non-linearly in the production of outputs. The performance analyses also demonstrated that, consequently, the response of the DSS under various conditions was not always easy to explain and the establishment of input-output relationships was not a straight-forward process. This, despite the fact that the structure of the DSS was relatively simple. This leads to the assessment that fuzzy systems may exhibit some complex behavior, even if they possess a simple structure. Therefore, response analysis constitutes an important step when developing fuzzy systems; analyses permit a better understanding of the interactions among rules, membership functions, input values and outputs.

Another aspect which needs to be considered when developing fuzzy systems is that the potential number of rules increases rapidly when combining variables with a large number of sets; three variables that are each represented by 5 fuzzy sets will lead to 125 possible rules. In this case, redundant rules (or rules with impossible outcomes) may need to be eliminated to ease the management of the RS and to accelerate the inference process. It must also be considered that, with fuzzy sets, many rules can be fired during an inference process since variables can take more than one value. It may be necessary, therefore, to keep track of many parallel reasoning lines, which requires additional computing and memory resources.

Future research plans include the development of DSS components concerning diagnostics and recommendations about other aspects of dairy cattle breeding (e.g., replacement, mating). Results obtained in this project constitute a basis for the elaboration of a methodology and of a complete framework that will help in that area. Research will be pursued on the methodological aspects. One particular aspect which will be investigated concerns the integration of the expertise of multiple experts for the development of a particular system. For the prototype DSS, the development was based on the combination of the expertise of three experts to form one set of rules and membership functions. This approach was based on the requirement that a consensus be reached among experts. However, during system development, there was not always complete agreement among experts. For example, there were variations about the exact meaning of certain fuzzy qualifiers and their numerical expression; this might have led to the establishment of different membership functions. A consensus was reached among experts on all other aspects but, in theory, different sets could have been used for the variables. Also, different rules could have been composed, and even different variables might have been used. In all these cases, instead of trying to reach a consensus so as to amalgamate all knowledge into one component, an alternative approach would be to aggregate various components in the DSS; components which would each contain one individual's expertise. With this last approach, final conclusions would be reached through conflict resolution. Research on this topic has already been initiated (9) and alternative methods for knowledge encoding might be used in future development.

Although the system which has been developed to date does not take account of all factors which a producer might consider when making culling decisions, it has shown promising results in preliminary studies in terms of agreement between decisions suggested by the experts and decisions suggested by the DSS. As it is, it can already be used for cow rating, without too much modification. Indeed, one of the interesting aspects observed during this project is that the degrees of membership associated with the descriptors of the fuzzy variable 'culling' could be used for ranking of cows within a herd. Such an approach might constitute an interesting alternative to rating methods currently used by some DHIs. This may even constitute an approach which would be preferable for many producers, who would themselves make the final decisions about culling the cows or not; they would use the ranks produced by the DSS in conjunction with other factors to base their final decision.

CONCLUSION

Fuzzy logic permits the encoding of knowledge using a terminology which is close to that used by experts, i.e., based on linguistic descriptions. This eases the encoding of knowledge, which may accelerate the development and the implementation of DSSs that accurately reproduce the reasoning of specialists. This would make them more rapidly available to producers. Since fuzzy logic seems promising in the development of knowledge-based systems, it will be used to develop a more complete DSS aimed at helping dairy producers to establish their herd breeding policy, their herd breeding programme and their breeding programme for individual cows.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada and the Cattle Breeding Research Council of Canada. Data were provided by the Québec Dairy Herd Analysis Service (Programme d'analyse des troupeaux laitiers du Québec).

REFERENCES

1. Ambuel, J.R., T.S. Colvin and D.L. Karlen. 1994. A fuzzy logic yield simulator for prescription farming. Transactions of the ASAE 37(6): 1999-2009.

2. Edwards, D.B. and J.R. Canning. 1995. Developing a radio-controlled log skidder with fuzzy logic autonomous control. Transactions of the ASAE 38(1):243-248.

3. Graham, I. and P.L. Jones. 1988. Expert systems: knowledge, uncertainty and decision. Chapman and Hall, New York, USA. 363 pp.

4. Grinspan, P., Y. Edan, H.E. Kahn and E. Maltz. 1994. A fuzzy logic expert system for dairy cow transfer between feeding groups. Transactions of the ASAE, 37(5):1647-1654.

5. Heatwole, C.D. and T.L. Zhang. 1990. Representing uncertainty in knowledge-based systems: confirmation, probability, and fuzzy set theories. Transactions of the ASAE 33(1):314-323.

6. Holsapple, C.W., and A.B. Whinston. 1986. Manager's guide to expert systems using GURU. Dow Jones-Irwin, Homewood, Illinois, USA. 312 pp.

7. Leung, K.S. and W. Lam. 1988. Fuzzy concepts in expert systems. Computer 21(9):43-56.

8. Micro Data Base Systems. 1991. GURU - Integrated components manual. Micro Data Base Systems, Inc., Lafayette, Indiana, USA.

9. Strasser, M., R. Lacroix, R. Kok and K.M. Wade. 1995. Knowledge engineering for fuzzy decision-support systems. CSAE Paper no. 95-612.

10. Thangavadivelu, S. and T. S. Clovin. 1991. Decision analysis using linguistic approximation in an agricultural environment. p 98-103 *in* Proceedings of the Simulation Multiconference on Artificial Intelligence and Simulation. Simulation Series 23(4).

11. Williams, T. 1992. Fuzzy logic is anything but fuzzy. Computer design, 31(4):113-127.

12. Zadeh, L.A. 1989. Knowledge representation in fuzzy logic. IEEE Transactions on Knowledge and Data Engineering 1(1):89-100.

13. Zimmermann, H.-J. 1991. Fuzzy set theory and its applications (2nd edition). Kluwer Academic Publishers, Norwell, Massachusetts, USA. 399 pp.

Table 1.1 - Fuzzy sets used in the prototype decision-support system.

VARIABLE NAME*	FUZZY DESCRIPTORS
Calving interval (CalvInt)	Short, Average, Long
Days to first breeding (FirstSrv)	Early, Average, Late
Number of breedings (NumServ)	Few, Average, Many
Reproductive efficiency (ReprEff)	Unsatisfactory, Poor, Good, Excellent
Production index (ProdIdx)	Very low, Low, Medium, High, Very high
Culling (Cull)	Yes, No

*The name indicated between brackets is used in appendices 1.1 to 1.3.

Table 1.2 - Inference results for Herd 9.

Cow identification	Calving interval [d]	Days to first breeding	Number of breedings	Lactation number	Production index [kg]	DM to 'yes'	DM to 'no'	Final culling decision
448	448	108	3.0	2	15838	0	98	no
453	412	95	1.7	3	15616	9	95	no
464	410	99	1.7	3	16357	0	92	no
473	395	93	1.5	2	17983	0	92	no
494	373	90	1.0	2	15914	0	83	no
420	466	87	2.5	3	17364	0	82	no
424	380	77	3.0	4	18278	0	82	no
418	381	80	1.7	4	17323	24	81	no
486	435	74	2.0	2	15432	18	81	no
488	393	84	1.5	2	15448	18	81	no
429	397	96	1.3	4	16214	55	77	no
430	393	71	1.5	4	15358	22	77	no
454	403	110	2.0	2	15356	22	77	no
474	362	81	1.0	2	15593	10	76	no
395	436	85	2.3	5	18629	0	75	no
465	352	84	1.0	3	18848	0	75	no
467	366	71	2.5	2	16809	0	75	no
490	405	84	1.5	2	15012	39	72	no
487	362	89	1.0	2	15007	40	59	unknown
431	416	103	1.3	4	16194	86	55	yes
384	414	68	2.3	5	15870	47	53	unknown
461	438	78	1.7	3	14234	78	34	yes
422	410	93	1.8	5	14197	87	19	yes
492	390	86	2.0	2	13981	91	15	yes
458	363	66	1.7	3	13669	93	0	yes
489	423	99	1.0	2	13379	83	0	yes
485	373	89	3.0	2	12981	75	0	yes

Table 1.3 - Inference results for Herd 21.

Cow identification	Calving interval [d]	Days to first breeding	Number of breedings	Lactation number	Production index [kg]	DM to 'yes'	DM to 'no'	Final culling decision
412	411	76	2.3	4	20171	0	94	no
430	388	86	2.3	3	18274	0	90	no
440	393	70	2.0	3	17991	3	89	no
450	359	77	3.0	2	20305	0	89	no
439	364	79	1.0	3	19159	0	85	no
453	447	70	2.5	2	20446	0	84	no
442	368	86	2.0	3	19167	0	83	no
355	411	90	1.8	6	20521	0	82	no
432	393	75	2.3	3	18704	0	81	no
434	430	75	2.3	3	20591	0	80	no
461	367	88	2.0	2	17706	17	77	no
427	431	98	2.3	3	18846	0	76	no
428	540	85	3.0	2	20795	0	76	no
426	369	72	1.7	3	17256	39	60	no
459	381	131	1.5	2	17193	42	57	unknown
454	357	83	2.5	2	17187	43	56	unknown
451	388	94	1.0	2	16503	77	22	yes
455	349	58	2.0	2	12974	100	0	yes
417	368	87	1.0	4	15297	76	0	yes
452	388	96	1.0	2	14847	76	0	yes
437	416	74	2.7	3	15181	75	0	yes

Table 1.4 - Inference results for Herd 30.

Cow identification	Calving interval [d]	Days to first breeding	Number of services	Lactation number	Production index [kg]	DM to 'yes'	DM to 'no'	Final culling decision
506	689	87	2.5	2	25849	0	100	no
465	516	173	2.0	2	23171	0	97	no
501	361	93	2.0	2	19448	0	89	no
428	511	183	1.8	7	21000	11	88	no
493	388	98	1.3	3	19798	0	84	no
508	400	115	1.0	2	20779	0	82	no
481	397	97	1.3	5	19120	5	77	no
451	529	124	3.5	6	21984	0	76	no
513	448	87	1.5	2	20002	0	76	no
421	418	102	1.6	5	20178	0	75	no
469	455	84	2.0	3	20134	0	75	no
520	402	74	2.5	2	17819	70	29	yes
453	407	83	2.2	6	17704	75	24	yes
489	503	126	1.5	2	17650	78	21	yes
527	394	77	3.0	2	17648	78	21	yes
434	493	148	2.2	5	19612	80	19	yes
529	436	125	2.0	2	13016	100	0	yes
526	455	32	2.5	2	17120	94	0	yes
532	416	131	1.0	2	17036	91	0	yes
475	656	165	3.8	6	18841	84	0	yes
517	476	71	3.5	2	15782	80	0	yes

Table 1.5 - Herds' average and standard deviation for calving interval, days to first breeding, number of breedings and production index.

HERD NUMBER	NUMBER OF COWS	CALVING INTERVAL [d]		DAYS TO FIRST BREEDING [d]		NUMBER OF BREEDINGS [d]		PRODUCTION INDEX [kg]	
		Avg.	(Std.)	Avg.	(Std.)	Avg.	(Std.)	Avg.	(Std.)
1	21	393	31	78	13	1.8	0.6	16288	1961
2	20	379	90	73	11	1.5	0.6	14812	2251
3	19	422	86	92	19	1.1	0.2	15485	1851
4	18	405	53	75	22	2.4	0.9	14580	1812
5	29	384	27	71	10	1.9	0.6	14189	1645
6	62	389	38	83	19	1.5	0.5	13524	1555
7	22	431	62	90	28	1.8	0.6	13148	1652
8	19	407	45	78	12	1.9	0.6	15252	1745
9	27	400	29	87	12	1.8	0.6	15810	1560
10	19	404	45	87	18	1.7	0.7	12856	1529
11	25	386	32	84	17	1.9	0.8	17420	1995
12	13	369	22	78	17	1.1	0.2	14859	1652
13	7	427	35	115	24	1.8	0.6	20341	1829
14	24	417	57	81	15	2.5	1.0	14892	2147
15	18	381	19	80	7	1.8	0.6	13499	1877
16	22	404	44	91	17	2.1	0.9	14324	1793
17	18	396	23	82	14	1.9	0.6	15027	1622
18	16	393	47	80	12	2.6	1.5	17017	3134
19	14	422	92	101	50	1.2	0.3	13911	2254
20	19	416	59	109	28	1.8	0.8	16947	2768
21	21	396	43	83	15	2.0	0.6	18053	2185
22	24	401	36	90	15	1.9	0.4	13859	1928
23	17	390	32	83	12	2.2	0.8	15693	2526
24	29	414	39	94	19	1.8	0.6	16590	1638
25	14	401	57	65	12	2.6	1.0	13984	985
26	12	418	62	93	24	1.7	0.4	13947	1620
27	33	398	46	82	11	2.0	1.1	14407	2161
28	49	394	53	83	22	1.4	0.5	14040	1996
29	21	400	36	77	9	2.3	1.0	16251	1626
30	21	464	84	108	37	2.1	0.8	19223	2688

Table 1.6 - Inference results for 30 herds.

HERD NUMBER	NUMBER OF COWS	NO		YES		UNKNOWN	
		Count	%	Count	%	Count	%
1	21	15	71	6	29	0	0
2	20	9	45	8	40	3	15
3	19	14	74	5	26	0	0
4	18	12	67	5	28	1	6
5	29	17	59	9	31	3	10
6	62	39	63	15	24	8	13
7	22	13	59	5	23	4	18
8	19	11	58	6	32	2	11
9	27	18	67	7	26	2	7
10	19	11	58	6	32	2	11
11	25	14	56	11	44	0	0
12	13	9	69	3	23	1	8
13	7	5	71	1	14	1	14
14	24	13	54	11	46	0	0
15	18	8	44	7	39	3	17
16	22	13	59	6	27	3	14
17	18	12	67	5	28	1	6
18	16	10	63	6	38	0	0
19	14	10	71	4	29	0	0
20	19	12	63	5	26	2	11
21	21	14	67	5	24	2	10
22	24	15	63	8	33	1	4
23	17	9	53	7	41	1	6
24	29	16	55	6	21	7	24
25	14	10	71	3	21	1	7
26	12	9	75	2	17	1	8
27	33	16	49	16	49	1	3
28	49	28	57	18	37	3	6
29	21	15	71	6	29	0	0
30	21	11	52	10	48	0	0

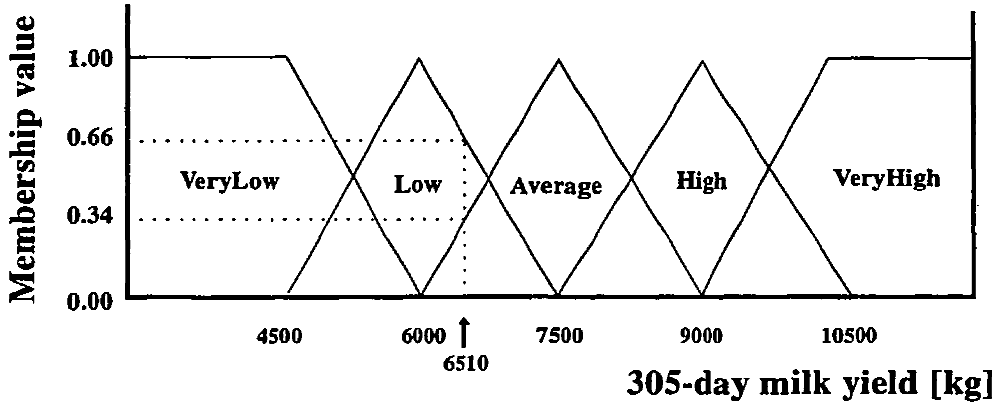

Figure 1.1 - Illustration of fuzzy sets characterizing 305-day milk yield.

RULE 1: **IF** MilkProd = Low **OR** ReprodEff = Low
THEN Culling = Yes

RULE 2: **IF** MilkProd = Average **AND** ReprodEff = VeryLow
THEN Culling = Yes

RULE 3: **IF** MilkProd = Average **AND** ReprodEff = Low
THEN Culling = No

Figure 1.2 - Example of a rule set involving fuzzy variables.

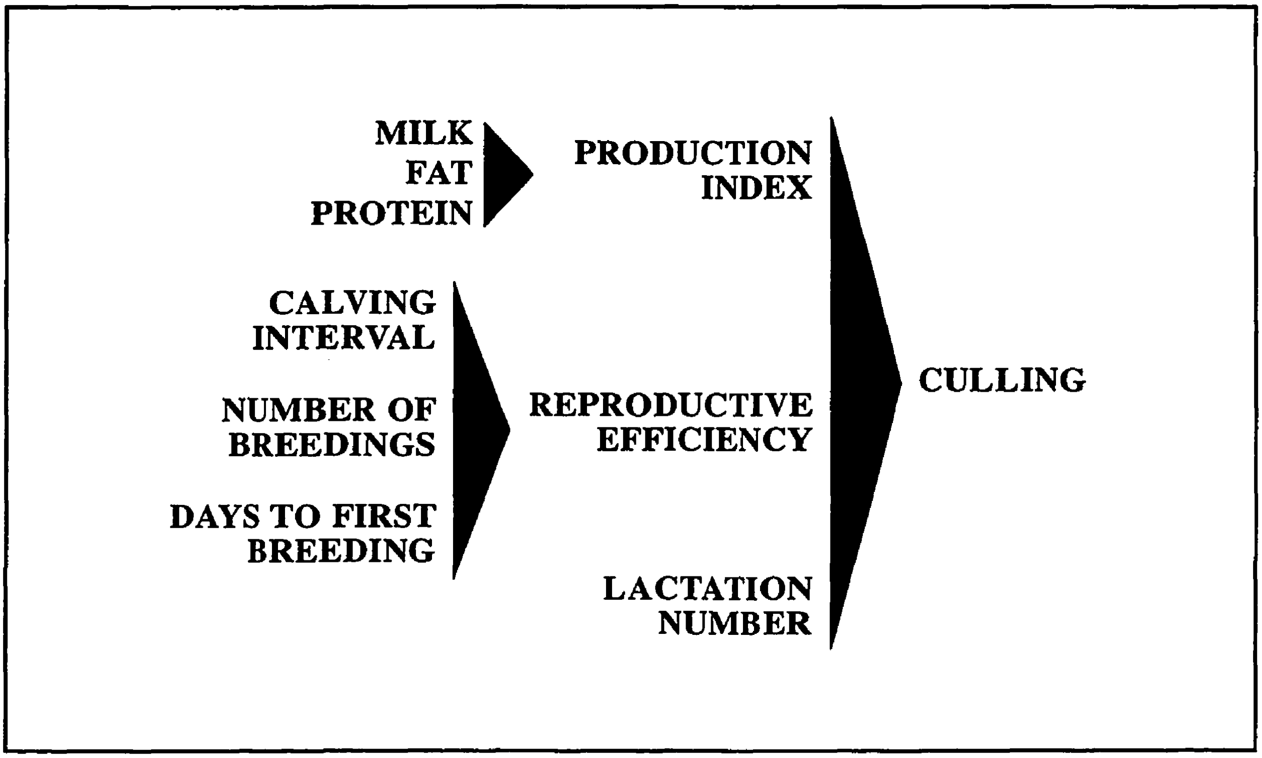

Figure 1.3 - Variables involved in the overall decision-making process.

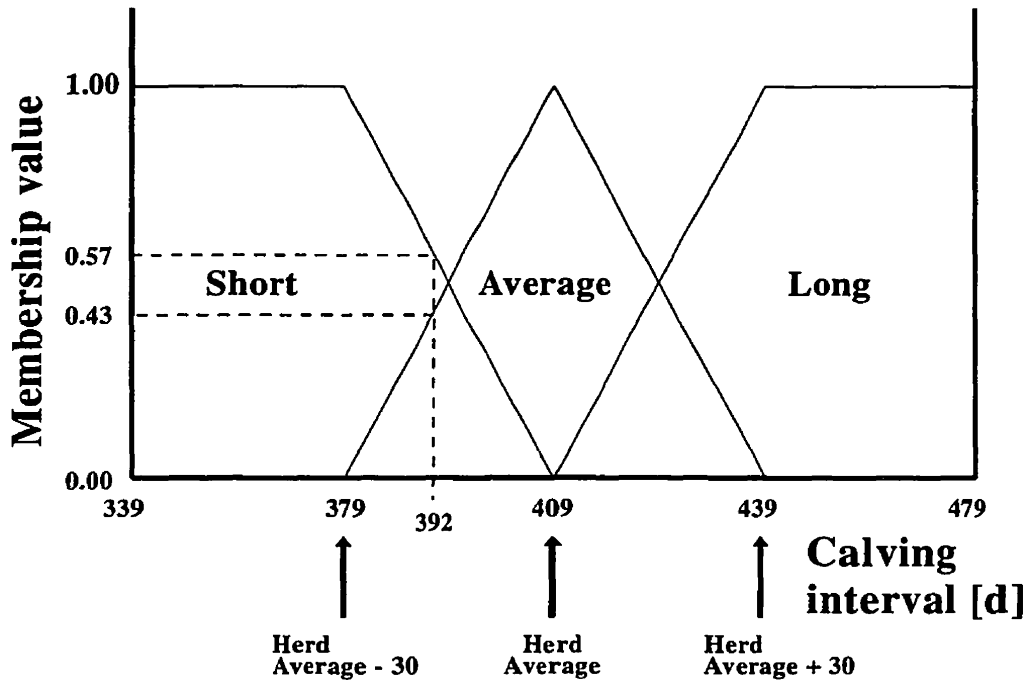

Figure 1.4 - Fuzzy sets for the variable 'calving interval'.

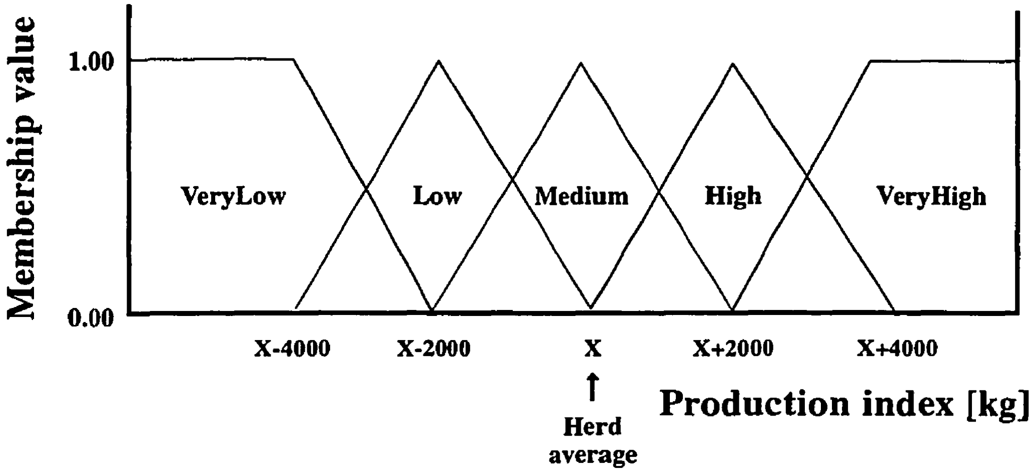

Figure 1.5 - Fuzzy sets for the variable 'production index'.

FROM 1st RULE SET

Input variables are:

1) Calving interval =
(average, .98), (short, .02)

2) Days to first breeding =
(late, .85), (average, .15)

3) Total number of breedings =
(average, .66), (many, .34)

RULE 5 (fired)
Calving interval is (short, .02), days to first breeding is (average, .15), and number of breedings is (average, .66):
Therefore reproductive efficiency is (good, .02).

RULE 6 (fired)
Calving interval is (short, .02), days to first breeding is (average, .15), and number of breedings are (many, .33):
Therefore reproductive efficiency is (good, .04).

RULE 8 (fired)
Calving interval is (short, .02), days to first breeding is (late, .85), and number of breedings are (average, .66):
Therefore reproductive efficiency is (poor, .02).

RULE 9 (fired)
Calving interval is (short, .02), days to first breeding is (late, .85), and number of breedings are (many, .34):
Therefore reproductive efficiency is (poor, .04).

RULE 14 (fired)
Calving interval is (average, .98), days to first breeding is (average, .15), and number of breedings are (average, .66):
Therefore reproductive efficiency is (good, .18).

RULE 15 (fired)
Calving interval is (average, .98), days to first breeding is (average, .15), and number of breedings are (many, .33):
Therefore reproductive efficiency is (poor, .18).

RULE 17 (fired)
Calving interval is (average, .98), days to first breeding is (late, .85), and number of breedings are (average, .66):
Therefore reproductive efficiency is (poor, .72).

RULE 18 (fired)
Calving interval is (average, .98), days to first breeding is (late, .85), and number of breedings are (many, .34):
Therefore reproductive efficiency is (unsatisfactory, .34).

FROM 2ND RULE SET

Input variables are:

1) reproductive efficiency =	(poor, .72), (unsatisfactory, .34), (good, .18)
2) production index =	(high, .76), (medium, .23)
3) lactation number =	9

RULE 2 (fired)
Production index is (high, .76):
Therefore culling is (no, .76).

RULE 4 (fired)
Production index is (medium, .23), and reproductive efficiency is (good, .18):
Therefore culling is (no, .80).

RULE 6 (fired)
Production index is (medium, .23), reproductive efficiency is (poor, .72), and lactation number is larger than 3:
Therefore culling is (yes, .23).

RULE 7 (fired)
Production index is (medium, .23) and reproductive efficiency is (unsatisfactory, .34):
Therefore culling is (yes, .41).

FROM 3RD RULE SET

The input variable is:

1) Culling =
(no, .80), (yes, .41)

RULE 3 (fired)
Since the difference between degrees of membership to NO and YES is more than .20, and the culling suggestion of the 2nd expert system is NO with a membership value larger than .60, the cow should not be culled.

Figure 1.6 - Example of an inference process.

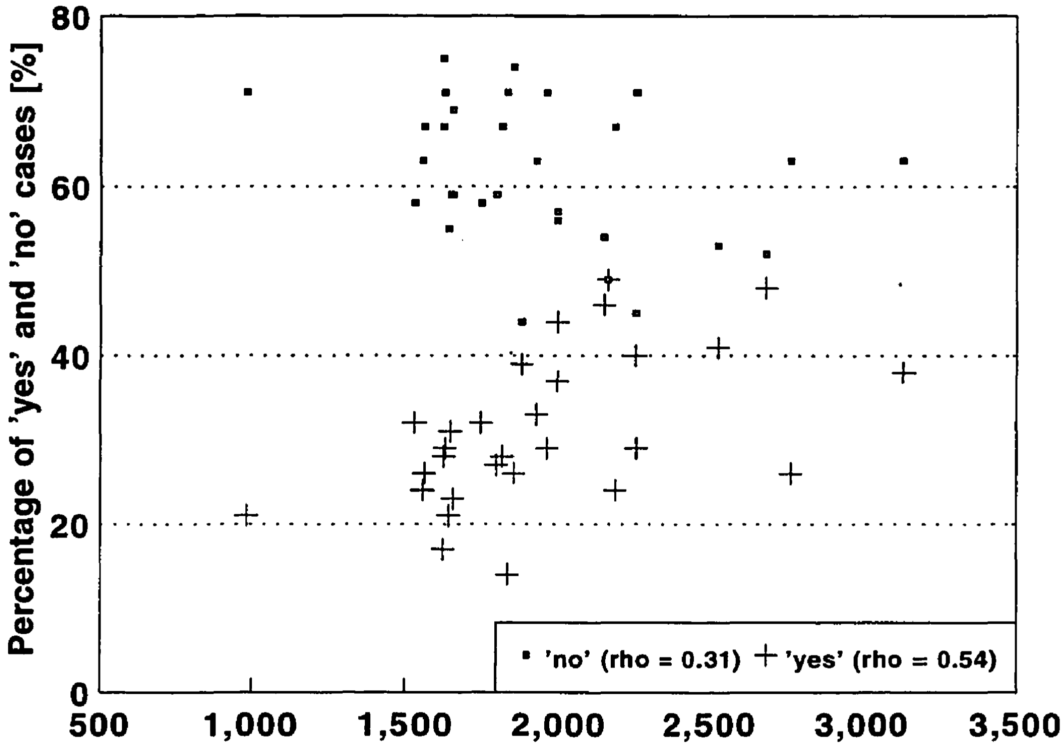

Figure 1.7 - Variation of the percentages of 'yes' and 'no' resulting from DSS's inferences, as a function of the herd standard deviation of the production index.

APPENDIX

Appendix 1.1 - First rule set (RS 1)

RULE 1:

IF: CalvInt = "Short" AND FirstSrv = "Early" AND NumServ = "Few"

THEN: ReprEff += "Excellent"

RULE 2:

IF: CalvInt = "Short" AND FirstSrv = "Early" AND NumServ = "Average"

THEN: ReprEff += "Good"

.....

RULE 7:

IF: CalvInt = "Short" AND FirstSrv = "Late" AND NumServ = "Few"

THEN: ReprEff += "Good"

RULE 8:

IF: CalvInt = "Short" AND FirstSrv = "Late" AND NumServ = "Average"

THEN: ReprEff += "Poor"

.....

RULE 13:

IF: CalvInt = "Average" AND FirstSrv = "Average" AND NumServ = "Few"

THEN: ReprEff += "Good"

RULE 14:

IF: CalvInt = "Average" AND FirstSrv = "Average" AND NumServ = "Average"

THEN: ReprEff += "Good"

.....

RULE 19:

IF: CalvInt = "Long" AND FirstSrv = "Early" AND NumServ = "Few"

THEN: ReprEff += "Poor"

RULE 20:

IF: CalvInt = "Long" AND FirstSrv = "Early" AND NumServ = "Average"

THEN: ReprEff += "Poor"

.....

RULE 25:

IF: CalvInt = "Long" AND FirstSrv = "Late" AND NumServ = "Few"

THEN: ReprEff += "Poor"

RULE 26:

IF: CalvInt = "Long" AND FirstSrv = "Late" AND NumServ = "Average"

THEN: ReprEff += "Unsatisfactory"

RULE 27:

IF: CalvInt = "Long" AND FirstSrv = "Late" AND NumServ = "Many"

THEN: ReprEff += "Unsatisfactory"

Appendix 1.2 - Second rule set (RS 2)

RULE 1:

IF: ProdIdx = "VHigh"

THEN: Cull += "No"

RULE 2:

IF: ProdIdx = "High"

THEN: Cull += "No"

RULE 3:

IF: ProdIdx = "Medium" AND ReprEff = "Excellent"

THEN: Cull += "No"

RULE 4:

IF: ProdIdx = "Medium" AND ReprEff = "Good"

THEN: Cull += "No"

RULE 5:

IF: ProdIdx = "Medium" AND ReprEff = "Poor" AND Lactation <= 3

THEN: Cull += "No"

RULE 6:

IF: ProdIdx = "Medium" AND ReprEff = "Poor" AND Lactation > 3

THEN: Cull += "Yes"

RULE 7:

IF: ProdIdx = "Medium" AND ReprEff = "Unsatisfactory" AND Lactation <= 3

THEN: Cull += "No"

RULE 8:

IF: ProdIdx = "Medium" AND ReprEff = "Unsatisfactory" AND Lactation > 3

THEN: Cull += "Yes"

RULE 9:

IF: ProdIdx = "Low"

THEN: Cull += "Yes"

RULE 10:

IF: ProdIdx = "VLow"

THEN: Cull += "Yes"

Appendix 1.3 - Third rule set (RS 3)

RULE 1:

IF: HICF(Cull)>=60 & DiffCF<=20*

THEN: Culling = "Unknown"

RULE 2:

IF: HICF(Cull)>=60 & DiffCF>20 & HIVAL(Cull)="Yes"

THEN: Culling = "Yes"

RULE 3:

IF: HICF(Cull)>=60 & DiffCF>20 & HIVAL(Cull)="No"

THEN: Culling = "No"

RULE 4:

IF: HICF(Cull)<60

THEN: Culling = "Unknown"

---

* HICF() is a function in GURU that returns the highest degree of membership for all values of a variable. The GURU function HIVAL() returns the value corresponding to the highest degree of membership. DiffCF() returns the difference between the highest and the lowest degree of membership for the variable "Cull".

CONNECTING STATEMENT

During the construction of the prototype DSS, many questions arose; examples included the establishment of membership functions, and the development of rule sets which would mimic the reasoning of the experts. Another example is related to the way in which knowledge is implemented when it is elicited from more than one expert. The number of questions which needed to be addressed was large enough to give rise to the need for the establishment of a more thorough methodology before constructing a second generation prototype. It was also felt that this methodology should be part of a more global framework which would help in determining the orientation taken when developing a fuzzy DSS, and which would assist in the development process. The framework needed to present possible decisions which need to be made, various alternatives, and other aspects which need to be taken into account.

CHAPTER 2

A Framework for the Development of Fuzzy Decision-Support Systems

M. STRASSER, R. LACROIX, R. KOK, and K.M. WADE

ABSTRACT

Because of the increased complexity of the modern farm, the agricultural sector can benefit through the introduction of decision-support systems. Such systems can be used to supplement current management practices by performing repetitive tasks and therefore allowing the producer to devote more time to other issues. The applicability of fuzzy-logic enhanced decision-support systems has been examined through the construction of a prototype fuzzy decision-support system for dairy cattle culling decisions. An analysis of the development process of this prototype demonstrated the importance of creating a thorough methodology for the elicitation and implementation of knowledge. From the analysis, it was also possible to observe the differences in how experts convey their knowledge. Because of its inherent subjectivity, experts communicate their knowledge in their respective personal languages. As a result, any decision-support system designed using an expert's personal language requires the creation of a translation process whereby the public language variables, considered by the expert, are translated into the experts' personal language and back again. Fuzzy mathematics can be used to this end. A framework for the development of fuzzy decision-support systems was established, consisting of four phases: the project groundwork phase, elicitation of knowledge from the expert, implementation of that knowledge, and system validation. In this framework, it is proposed that, in the case of multiple experts, knowledge can be amalgamated or aggregated. If the knowledge of multiple experts is aggregated, a conflict resolver must be developed ensuring that, after defuzzification, there exists only one single, crisp system output.

(**Key words**: dairy cattle breeding, decision-support systems, fuzzy logic, knowledge engineering)

INTRODUCTION

The modern dairy farmer is required to make many complex decisions on a continuous basis and any shortcoming in the decision-making process can quickly result in sub-optimal management practices (Hogeveen *et al.* 1994.). While the use of computer-based technologies has been forecast as being an effective supplement to existing management practices in the dairy sector (Tomaszewski, 1992), in many instances, it has simply resulted in an information "overload", especially at the farm level, where the information must be "distilled" before utilization. With this in mind, decision-support systems (DSS) have been proposed as an effective tool for delivering precise and specific advice (Greer *et al.*, 1994).

A research program is currently in progress with the objective of creating a global DSS for dairy cattle breeding. The DSS will be constructed using fuzzy-logic expert systems because of their ability to deal with heuristic knowledge, as well as the vagueness and uncertainty often associated with the reasoning processes of experts. As a first step towards the creation of this global DSS, research has focussed on the exploration of a narrow subject area, namely, dairy cattle culling decisions. A prototype DSS was constructed to learn about the various components and characteristics of fuzzy DSSs, and to become familiar with procedures for knowledge acquisition and implementation in a fuzzy-logic context (Lacroix *et al.*, 1994).

During the construction of the prototype DSS, many questions arose; examples included the establishment of membership functions, and the development of rule sets which would mimic the reasoning of the experts. Another example is related to the way in which knowledge is implemented when it is elicited from more than one expert. The number of questions which needed to be addressed was large enough to give rise to the need for the establishment of a more thorough methodology before constructing a second generation prototype. It was also felt that this methodology should be part of a more global framework which would help in determining the orientation taken when developing a fuzzy DSS, and which would assist in the development process. The framework needed to present possible decisions which need to be made, various alternatives, and other aspects which need to be taken into account.

A great deal of research has been conducted over the last decade related to the development process of knowledge-based systems, and there exists a great deal of documentation related to knowledge engineering techniques. For example, research has been conducted on the interview process (Greenwell, 1988; Jones *et al.*, 1987; Oltjen *et al.*, 1990; Spangler *et al.*, 1989; Wright and Ayton, 1987), domain modelling (Allemang and Rothenfluh, 1992; Compton *et al.*, 1993) and structuring of acquired knowledge (Ascough II *et al.*, 1990). Research has also been carried out to produce methods helping in the development of fuzzy (logic-based) controller expert systems (Wakami, 1991; Yen and Pfluger, 1991). However, there is a scarcity of information related to the techniques for the construction of fuzzy-logic based decision-support systems.

Because it was felt that a methodological framework was required to pursue the development of DSSs for dairy cattle breeding, and because a framework suiting our needs was not found, the decision was made to create one. The development of this framework was accomplished through an analysis of the prototype previously developed, in combination with existing literature. The framework encompasses the groundwork necessary at project outset, the knowledge elicitation and implementation processes, the defuzzification procedure, and the resolution of "conflict" when multiple-experts are involved. The first part of this paper is used to present a description and analysis of the prototype, while the second part concerns the knowledge engineering framework.

PROTOTYPE FUZZY EXPERT SYSTEM DEVELOPMENT

In order to gain better insight into fuzzy DSS development, the procedures used in the creation of the prototype were analysed. These procedures could be summarized by the five following steps: i) determination of the goal for the fuzzy DSS; ii) selection of the system variables and the linguistic labels of the fuzzy sets; iii) acquisition of the production rules (since it was decided that the fuzzy DSS would be rule-based); iv) determination of the membership functions; and v) validation of the prototype. The knowledge obtained for the fuzzy DSS was acquired using an interview approach with three experts in the domain of dairy cattle breeding. The resulting knowledge was combined to yield a single set of variables, fuzzy sets, production rules and membership functions. This process of combining the knowledge of more than one expert into a single system, is hereafter referred to as "*amalgamation*".

The first development objective was to determine the goal of the prototype. Informal discussions on dairy research needs with an academic from the local Department of Animal Science prompted us to focus our prototype on decisions regarding the culling of individual cows on a dairy farm. Once this was determined, two experts were recruited to participate in the project: another academic from the local Department of Animal Science and a representative of the local Dairy Herd Improvement Agency (DHIA). These experts were then consulted through the use of informal interviews. In order to facilitate rapid prototype development, only variables for which values were readily obtainable from the local DHIA were considered as potential fuzzy DSS inputs. The first expert was asked to provide linguistic labels for each of the variables. For example, the expert decided that the linguistic labels "few", "average", and "many" sufficiently characterized the variable "Number of Breedings". These linguistic labels were also used as the names of the fuzzy sets. The reader is referred to Table 2.1 for an example of the linguistic labels used in the prototype.

Once the variables and linguistic labels were obtained, the next objective was to determine the relationships among the variables. The creation of the production rules was structured with an IF...THEN... syntax, where the relationship between the antecedent, or the *IF* part of the rule, and the consequent propositions, or the *THEN* part of the rule, yields the rule's output. An example of a production rule, as illustrated in Figure 2.1, might be:

IF "Production" is "very low" **OR** "Reproductive Efficiency" is "bad"

THEN "Culling" is "yes"

The relationships for the prototype were obtained by presenting the experts with a list of variables and their linguistic labels and asking for the most appropriate conclusion. For example, given the variable "Production" (with an associated linguistic label of "very low"), and the variable "Reproductive Efficiency" (with an associated linguistic label of "bad"), the experts concluded that the cow should be culled. When agreement between experts was not exact, the knowledge engineer used personal judgement to combine or amalgamate the differences and to create the remaining expert system rules. This resolution was accomplished by taking an approximate average of all appropriate responses.

After the production rules were established, the membership functions were determined. Various methods of assessing these functions from data analysed by the experts were considered. However, to simplify development, the procedure which was finally retained consisted simply of drawing the functions from "critical point" values, with all fuzzy sets being represented by triangles and trapezoids. The critical points corresponded to the summits of thetriangles and the points of discontinuity in the top segment of the trapezoids. These values were assessed from information obtained from the experts during interviews. Since it was noted that the experts consistently based their culling decision process for a particular cow on the herd-average values for the various input variables, it was decided that these values would be used to define the central critical points. The other critical points were determined with the help of one expert, and validated with the two other experts. For example, in the case of the "calving interval" variable, the two other critical values were set at the herd average plus and minus 30 days, while the critical points for "days to first breeding" were set at herd average plus and minus 10 days. For a representation of sample calving interval fuzzy sets and the critical points, the reader is referred to Figure 2.2.

Once a sufficient knowledge-base was acquired, all the components (i.e., the production rules, membership functions, etc.), were integrated into a single decision-support system, which consisted of three rule-based expert systems (Figure 2.3). The first expert system combined the variables "Calving Interval", "Days to First Breeding", and "Total Number of Breedings" to produce a value for the variable "Reproductive Efficiency". A second expert system combined "Reproductive Efficiency", "Production Index", and "Lactation Number" -- a crisp (quantitative) variable -- to yield a culling decision which was fuzzy (i.e., "yes" or "no", with an associated membership value for each). It should be noted that the "Production Index" was arrived at by combining 305-day yields for milk, fat, and protein production, using economic weights similar to those currently in place for the Canadian dairy industry. A third expert system was created to defuzzify the culling decision, resulting in a crisp system output. A more detailed description of the prototype can be found in Lacroix *et al.* (1994), while a graphical representation of the decision-tree along with the organization of the expert systems used in the in the prototype are shown in Figure 2.3.

The development of the prototype allowed for a better understanding of the complexities related to the construction of fuzzy decision-support systems. Although the development process was, for the most part, free of difficulties, there are some points which should be mentioned. Concerning the participation of experts, there were occasional inconveniences with the scheduling and implementation of discussions. While this might be expected in such a procedure, its impact needs to be addressed, especially with regard to both the number of experts involved in such a project and the number of interviews required with each one. Another, perhaps, more significant concern, is the way in which experts differed in their interpretation of the system components (i.e., the variables considered, the production rules, the linguistic labels, and the membership functions). This demonstrated that experts have different personal definitions for the variables they use, how they relate them to each other, and how they define the fuzzy sets: experts convey their knowledge in their respective personal languages. For example, it is common to hear dairy producers describe a cow's milk production in terms of "high", "medium", or "low" which conveys a limited amount of information to others, depending on their own personal languages. In order for this personal language to be understood, it must be translated to public language. Public language refers to language used as a common reference by all members of a community. As an example of public language, a producer who states that a cow produced 10,500 kg. of milk would be understood by all. It is also possible that, in defining variables, various experts may use different linguistic labels in referring to the same membership functions, or conversely, they may use the same linguistic labels to describe different membership functions. Therefore, by amalgamating the opinions of multiple experts, as was the case with the prototype, the original contribution of the individual expert decreases and the knowledge-base becomes diluted. Because of this, it may be advantageous to use an approach in which the knowledge of each expert is kept separate. Such an approach is hereafter called "*aggregation*" and constitutes an alternative to *amalgamation.*

A FRAMEWORK FOR FUZZY DSS DEVELOPMENT

Generalities

In general terms, fuzzy DSS development can be characterized by four distinct stages: groundwork, knowledge elicitation, knowledge implementation, and system validation. The groundwork phase includes all work prior to knowledge elicitation and serves to prepare the foundation upon which the fuzzy DSS will be constructed. Knowledge elicitation is the knowledge transfer process from a source (e.g., an expert) to an intermediary (e.g., a knowledge engineer). Knowledge implementation is the process of interpreting, organizing, and encoding the elicited knowledge so that the resultant knowledge-base is representative of the experts' reasoning processes. Finally, validation is the process of fine-tuning the system and determining its accuracy and relevancy. Together these four stages form a skeleton for fuzzy DSS development.

Groundwork

The primary objective of the groundwork stage is to determine the goal of the fuzzy DSS. In selecting this goal one should consider research priorities, allowable project time, available expertise, and the final-product user. Once the goal has been established the knowledge engineer should perform the necessary research, referred to as pre-training, in order to obtain a basic understanding of the domain considered and to become familiar with the subject vocabulary (Jones, 1989; Jones *et al.*, 1987; Spangler *et al.*, 1989). There are different pre-training methods which can be employed, either on their own or in combination; consultation of the literature and tutorial sessions provided to the knowledge engineer by a project expert are two such methods. The latter allows for the establishment of a personal rapport between the project expert and the knowledge engineer which may be of benefit to the overall project development process (Jones, 1989; Spangler *et al.*, 1989). It should be noted that the term "project expert" refers to an expert involved as a consultant to the knowledge engineer, as opposed to a "domain expert" whose function is to supply the knowledge which is incorporated into the fuzzy DSS.

Obviously, the quality of the knowledge base is dependent upon the quality of its source. Accordingly, selection of the appropriate domain expert(s) is an important consideration in fuzzy DSS development. Choice of the domain expert(s) is often obvious; for example, the employee with the most experience. However, in some domains the choice may not be so clear. In the area of dairy cattle breeding one could choose among different types of domain experts; for example, researchers, dairy producers and dairy herd improvement agency (DHI) advisors. The first domain expert-type, the researchers, generally have strong theoretical knowledge about the domain but may have limited experience at the domain application level, while dairy producers have application level knowledge, but may have limited theoretical knowledge. The third domain expert-type, the DHI advisors, may have more theoretical knowledge than the producers and more applied knowledge than the researchers. Depending upon the project objectives it may also be advantageous to consider the use of multiple domain experts, due to differences which may exist in opinions as well as perspective. It should be noted that because of the potential for inter-expert knowledge variation, as well as a potential increase in development time, the inclusion of more than one domain expert should only be considered if it is anticipated that this will lead to a net gain in system performance.

The next stage, after selection of domain experts, involves determining the techniques which will be employed for the elicitation of expert knowledge. Many techniques are available and have been widely discussed in the literature, including questionnaires, interviews, and walkthroughs (Ascough *et al.*, 1990; Greenwell, 1988; Jones, 1989; Jones *et al.*, 1987; Oltjen *et al.*, 1990; Spangler *et al.*, 1989; Wright, 1987). While the technique chosen is dependent upon the project domain and the domain experts, the most commonly used technique is the interview. Other techniques are used if, for example, the expert is located too far away to conduct interviews and, in such instances, questionnaires may be more appropriate.

Knowledge Elicitation

After completion of the groundwork stage, the focus of the project shifts to the actual transfer of expertise, specifically, system variables, linguistic labels, production rules, and membership functions. Knowledge elicitation begins with the introduction of the project to the experts, choice of system variables, their definitions, and the linguistic labels of the associated fuzzy sets. Subsequent sessions are concerned with extraction of the relationships among the variables in order to define the production rule-sets. This process is followed by solicitation of the membership functions which translate public language input variables into the experts' personal languages and back again. Together, these sessions form the basis for the elicitation of expert knowledge for fuzzy DSS development.

Collection of the system variables, clarification of their respective definitions, and assignment of linguistic labels to the fuzzy sets can be accomplished using a number of procedures; a list of system variables and linguistic labels can be explicitly requested from the experts, or they can be derived from the interview transcripts. The first procedure has the advantage of allowing for *direct* acquisition, but might force the expert into providing "textbook" answers which may or may not reflect his/her actual decision-making process. The second procedure has the advantage that system variables and linguistic labels are less "prompted". However, they may not all be extracted during the interview. For obvious reasons, it is essential that the knowledge engineer obtain all significant system variables and linguistic labels during the interviewing process. Therefore, both techniques may need to be employed in order to obtain a complete and accurate set.

The production rules are expressed in terms of the system variables provided by the expert(s). These rules can be obtained through capture of the relationships among variables, as implicitly supplied by the experts, or they can be obtained explicitly by, for example, asking the experts to comment on all possible variable combinations. In order to provide an example of the latter process, the procedure used to obtain the production rules for the prototype's first expert system will be examined. The goal of this expert system was to determine the value of "Reproductive Efficiency" for a particular cow. The variables considered included "Calving Interval", "Days to First Breeding", and "Number of Breedings". "Reproductive Efficiency" was described with the linguistic labels "excellent", "good", "poor", and "unsatisfactory". "Calving Interval" was described with the linguistic labels "short", "average", or "long"; "Days to First Breeding" was described with "early", "average" or "late", and "Number of Breedings" was described with "few", "average", or "many". In determining the production rules, the experts were provided with a list of all possible combinations of variables and asked to state the results for each. For example, a "Calving Interval" of "short", combined with "Days to First Breeding" of "early", and "Number of Breedings" of "few" resulted in a "Reproductive Efficiency" of "excellent". Since there were three linguistic labels for each of the input variables, their combination resulted in a maximum of 27 possible rules. However, there is often overlap or redundancy in rules (e.g., "Reproductive Efficiency" may be deemed "excellent" for a "Calving Interval" of *either* "short" *or* "average", given the same two outcomes in the other variables) and so, expert systems can generally be streamlined by combining overlapping rules and removing redundant ones.

Subsequent to the acquisition of system variables and production rules, the membership functions of the fuzzy sets must be developed. They are used to correlate public language with the personal languages of the experts. As observed with the prototype fuzzy DSS, experts sometimes consider the same linguistic labels but, because they communicate in their respective personal language, they define them differently. For example, most experts would agree that very low production would be a reason to cull a cow, but these experts might differ in how they define "very low". One expert might qualify a production level of 3,000 kg. as "very low", whereas another would suggest that 4,000 kg. is "very low". The membership functions are used to translate from the crisp DSS input variables in public language to the fuzzy variables in personal language (i.e., for *fuzzification*), and vice-versa (i.e., for *defuzzification*). This is necessary if the production rules are to be developed using the personal languages of the experts. In Figure 2.4 the translation process for 305-day milk production is demonstrated. It can be seen that five fuzzy sets are used to cover the range for this variable. All values below 4,000 kg. are represented by the fuzzy set "very low" with a degree of membership of 1.0. Between 4,000 kg. and 5,500 kg. the degree of membership in the fuzzy set "very low" decreases while the degree of membership in the fuzzy set "low" increases. For example, a 305-day milk production value of 5000 kg. is represented by both the fuzzy set "very low" with a membership value of 0.35 (this is where the 5,000 kg. intersects the "very low" membership function), and the fuzzy set "low" with a membership value of 0.65.

In order to determine the membership functions, domain experts are first asked for the units of measure and the allowable range of values for a particular system variable. For example, the variable "Calving Interval" might be described using the units of measure "days" and might have a possible range of real values of "300 to 700". The next step involves the determination of the degree of association of a variable's value to a linguistic label of a fuzzy set, i.e., their degree of membership in the set. Three possible procedures are presented here. The first procedure, illustrated in Figure 2.5, consists of providing the domain expert with various values from the allowable range and asking him/her to associate a linguistic label with each value. For example, given a calving interval of 405 days, the domain expert might associate this value with the fuzzy set "average". The knowledge engineer would question the domain expert until enough values are obtained to cover the range of possible values. Apart from establishing the membership functions, these results could be used to visualize any trends. A second procedure, represented by Figure 2.6, is to present the domain expert with various values but, in this case, ask him/her to select two linguistic labels (and respective weightings) which describe each numerical value. Using our example of a 405-day calving interval, the domain expert might select '0.8 average' and '0.2 short' (i.e., the variable's value has a degree of membership of 0.8 out of 1.0 in the fuzzy set "average" and a degree of membership of 0.2 out of 1.0 in the fuzzy set "short"). A third procedure, illustrated in Figure 2.7, involves the introduction of a list of hedges (e.g., always, usually, most often, often, sometimes, rarely, never, etc.). In this procedure, the domain expert is first presented with a list of numerical values, the list of hedges, and the linguistic labels of the variable; the expert must choose the label which best represents each value, as well as the hedge which best describes the value's association to that label. For example, a 385-day calving interval might be described as "most often short" and a 365-day calving interval as "always short", etc. "Short" here represents the fuzzy set, while "always" represents the hedge. The domain expert could be prompted with as many cases as necessary to cover the entire range of the variable. Then, the list of hedges alone is presented to the domain expert in random order, and he/she is asked to provide a value, between 0 and 1, which expresses his/her confidence that the hedge represents the likelihood of an occurrence. For example, given the hedge "always", most experts would be 100% confident of an occurrence and would provide a value of 1 whereas a hedge of "rarely" might describe a value in the range of 2-5%. The hedges would then represent a particular membership value in a fuzzy set. Whereas the first procedure which was presented here requires that the knowledge engineer interpolate the results to determine the shape of the fuzzy sets and their representative equations, the data obtained from the last two procedures provides the knowledge engineer with an actual membership value for each value of a variable.

Knowledge Implementation

Once sufficient knowledge has been elicited, consideration must be given to its implementation. If the fuzzy DSS knowledge-base is derived from only one expert, implementation is straightforward. The system components (i.e., variables, fuzzy sets, membership functions, and rule-sets) are all created in the personal language of that expert and the DSS output, after defuzzification, is in public language. The situation becomes more complex if the knowledge of multiple domain experts is incorporated. In this case, the implementation approach needs to be chosen. Two options are available: the knowledge can be amalgamated, (i.e., combined into one common expert system), or it can be aggregated (i.e., kept apart in different expert systems). Various intermediary methods could also be envisaged but, in this framework, only the two main options are presented.

An amalgamated fuzzy DSS can be created in cases where the domain experts agree upon a common language and use this common (personal) language to convey their knowledge. Amalgamated fuzzy DSSs can also be created by the knowledge engineer by using personal judgement to combine the knowledge of the individual domain experts (a project expert might also be consulted for this). Amalgamation by consensus has the advantage over amalgamation carried out by the knowledge engineer in that the inter-expert variations in knowledge and experience are combined by the experts themselves and, therefore, the errors resulting from knowledge misinterpretation should be minimized. The amalgamation approach, which is illustrated in Figure 2.8, results in a single set of system variables, linguistic labels, membership functions and production rules (as developed for the prototype fuzzy DSS). Therefore, knowledge implementation following this approach leads to a DSS which is quite similar to that obtained with a single expert. The DSS output is unique and, after defuzzification, is in public language. One potential shortcoming of amalgamation is that it might not represent the decision-making process of any individual expert; it is a hybrid of the various personal languages of the domain experts, and may, in fact, result in a system that represents none of the experts sufficiently well to achieve the original goal. Validation must be performed to ensure that this does not occur.

A multiple-expert DSS can also be constructed by aggregating the domain experts' knowledge. In this approach, modules are developed for each domain expert, as illustrated in Figure 2.9. Each module consists of production rules, linguistic labels and membership functions constructed in the individual personal language of one expert, and the consultation of each module results in an individual decision. This decision will be in public language. Since it is unlikely that the outputs of the various modules will be exactly the same, a conflict resolver is required to combine the individual decisions into a unique DSS output. Many approaches can be envisaged for conflict resolution. For example, a conflict resolver could be constructed so as to exploit the differences existing between the expertise of each expert involved in the DSS development. Each expert brings his/her own experience and perspective to the project, and the resolver could be constructed so as to offer the end-user an opportunity to select the perspective (i.e., the module) with which he/she identifies, or which best reflects his/her own managerial priorities. An alternative approach would be that all modules receive the same weight, and that the system be designed so that the multiple module outputs are combined, resulting in a single system output. In this case, the simplest conflict resolver might use "voting" to select the most popular decision (e.g., if three modules return a culling value of "yes" and two modules respond with "no", the overall system output would be "yes"). Another possibility consists of leaving the responsibility of resolving conflict with the end-user. The knowledge engineer could also design a resolver which would attribute a different "weight" to each module, and all module decisions would be provided to the user who could then make his/her own final decision. Other techniques for the resolution of inter-expert conflict are possible and further research is needed in this area. It is important to note that conflict resolution can only be performed after defuzzification since, prior to defuzzification, the system outputs are in the experts' personal languages. It should also be noted that an aggregated fuzzy DSS might require a higher level of management than an amalgamated one due to the greater number of modules and the necessity of keeping track of multiple personal languages simultaneously.

Knowledge Validation

After an initial DSS has been assembled, field testing is necessary. In this, the knowledge base can be tested in its entirety, focussing on the results, or each component can be individually analysed. The system can be presented to the expert(s) to determine whether or not the system matches his/her/their judgement. If the DSS varies greatly from the expert(s)' decisions, other iterations through the knowledge elicitation and implementation cycle become necessary. If the expert(s) is (are) sufficiently satisfied with the results, the knowledge engineer finalizes DSS development through the construction of the system support routines and the interface. It is to be noted that validating a DSS which is constructed from multiple experts will be considerably more tedious than when only one expert is used. Again, this reinforces the suggestion that more than one expert should be used only when the benefits at least compensate for the inconveniences produced with this approach.

CONCLUSIONS

Through prototype development, it has been demonstrated that domain experts differ in their use of language and this must be considered when developing decision-support systems. If the knowledge of more than one domain experts is acquired for implementation into a fuzzy DSS, the knowledge can be amalgamated or aggregated. The latter approach ensures that the knowledge of each expert is maintained independently in his/her own personal language and would seem to allow more flexibility with regard to end-user applications. This approach will be investigated in more detail during the construction of a second generation prototype DSS for culling decisions.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada, Semex Canada and the Cattle Breeding Research Council of Canada. Data were provided by the Québec Dairy Herd Analysis Service (Programme d'analyse des troupeaux laitiers du Québec).

REFERENCES

Allemang, D., and T.E. Rothenfluh. 1992. Acquiring knowledge of knowledge acquisition: A self study of generic tasks. In Wetter, T., Althoff, K.-D., Boose, J., Gaines, B.R., Linster, M., and Schmalhofer, F. (Eds.). *Current Developments in Knowledge Acquisition-European Knowledge Acquisition Workshop 92*, pp. 353-372. Springer-Verlag.

Ascough II, J.C., L.J. Lane, and L. Deer-Ascough. 1990. Interactive knowledge acquisition and expertise transfer: overview and methodologies. American Society of Agricultural Engineers. Paper No. 90-755.

Compton, P., B. Kang, P. Preston, and M. Mulholland. 1993. Knowledge acquisition without analysis. In Aussenac, N., Boy, G., Gaines, B, Linster, J.-G., and Kodratoff, Y. (Eds.) *Knowledge Acquisition for Knowledge-Based Systems-European Knowledge Acquisition Workshop 93*, pp. 277-299. Springer-Verlag.

Edan, Y., P. Grinspan, E. Maltz, and H. Kahn. 1992. Fuzzy logic applications in the dairy industry. American Society of Agricultural Engineers. Paper No. 92-3600.

Greenwell, M. 1988. Knowledge Engineering for Expert Systems. Ellis Horwood Ltd. Chichester.

Greer, J.E., S. Falk, K.J. Greer, and M.J. Bentham. 1994. Explaining and justifying recommendations in an agricultural decision support system. Computers and Electronics in Agriculture, 11(1994). Elsevier Science Publishers B.V., Amsterdam. p. 195-214.

Hogeveen, H., M.A. Varner, D.S. Brée, D.E. Dill, E.N. Noordhuizen-Stassen, and A. Brand. 1994. Knowledge representation methods for dairy decision support systems. Journal of Dairy Science. 77:3704-3715.

Jones, J.W., P. Jones, and P.A. Everett. 1987. Combining expert systems and agricultural models: A case study. Transactions of the American Society of Agricultural Engineers. p. 1308.

Jones, P.H. 1989. Knowledge acquisition. *Knowledge Engineering in Agriculture*. ed. Barrett, J.R., and D.D. Jones. American Society of Agricultural Engineers. St. Joseph, Mich. p. 33-46.

Lacroix, R., M. Strasser, K. Wade, and R. Kok. 1994. A fuzzy-logic based decision-support system for dairy cattle breeding. American Society of Agricultural Engineers. Paper No. 94-3555.

Oltjen, J.W., G.E. Selk, L.G. Burditt, and R.E. Plant. 1990. Integrated expert system for culling management of beef cows. Computers and Electronics in Agriculture, 4(1990). Elsevier Science Publishers B.V., Amsterdam. p. 333-341.

Spangler, A.M., C.D. Ray, and K. Hamaker. 1989. Knowledge acquisition for expert system development. Computers and Electronics in Agriculture, 4(1989). Elsevier Science Publishers B.V., Amsterdam. p. 23-32.

Tomaszewski, M.A. 1992. Using advanced computer technologies to increase extension effectiveness. Journal of Dairy Science. 75:3242-3245.

Wakami, N. 1991. Fuzzy control and neural networks: Applications for consumer products. In Driankov, D., P. Eklund, A. L. Ralescu (Eds.) *Fuzzy Logic and Fuzzy Control. International Joint Congress on Artificial Intelligence 91*, pp. 99-105. Springer-Verlag.

Wright, G. and P. Ayton. 1987. Eliciting and modelling expert knowledge. Decision Support Systems, 3: 13-26.

Yen, J. and N. Pfluger. 1991. Using fuzzy logic in a mobile robot path controller. In Driankov, D., P. Eklund, A. L. Ralescu (Eds.) *Fuzzy Logic and Fuzzy Control. International Joint Congress on Artificial Intelligence 91*, pp. 135-142. Springer-Verlag.

Table 2.1 - Linguistic labels used in the development of the prototype decision-support system.

VARIABLE	LINGUISTIC LABELS				
Calving Interval	Short		Average		Long
Days to First Breeding	Early		Average		Late
Number of Breedings	Few		Average		Many
Production	Very High	High	Medium	Low	Very Low
Reproductive Efficiency	Excellent		Good	Poor	Unsatisfactory
Culling		Yes		No	

Sample Rule-Set

IF Production = "Medium" **AND** Reproductive Efficiciency = "Poor"	**THEN** Cull = "Yes"
IF Production = "Medium" **AND** Reproductive Efficiency = "Excellent"	**THEN** Cull = "No"
IF Production = "Medium" **AND** Reproductive Efficiency = "Good"	**THEN** Cull = "No"
IF Production = "Low" **OR** Production = "Very Low"	**THEN** Cull = "Yes"

Figure 2.1 - Sample rule set from the prototype decision-support system.

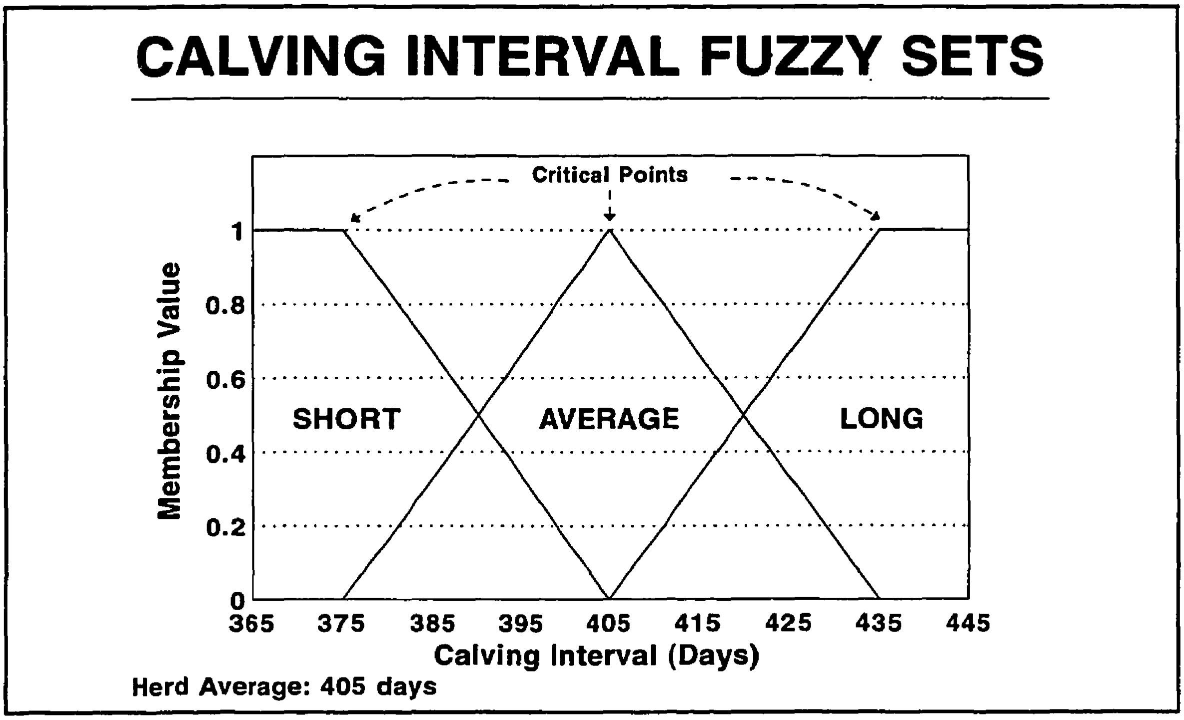

Figure 2.2 - Sample fuzzy sets with critical points.

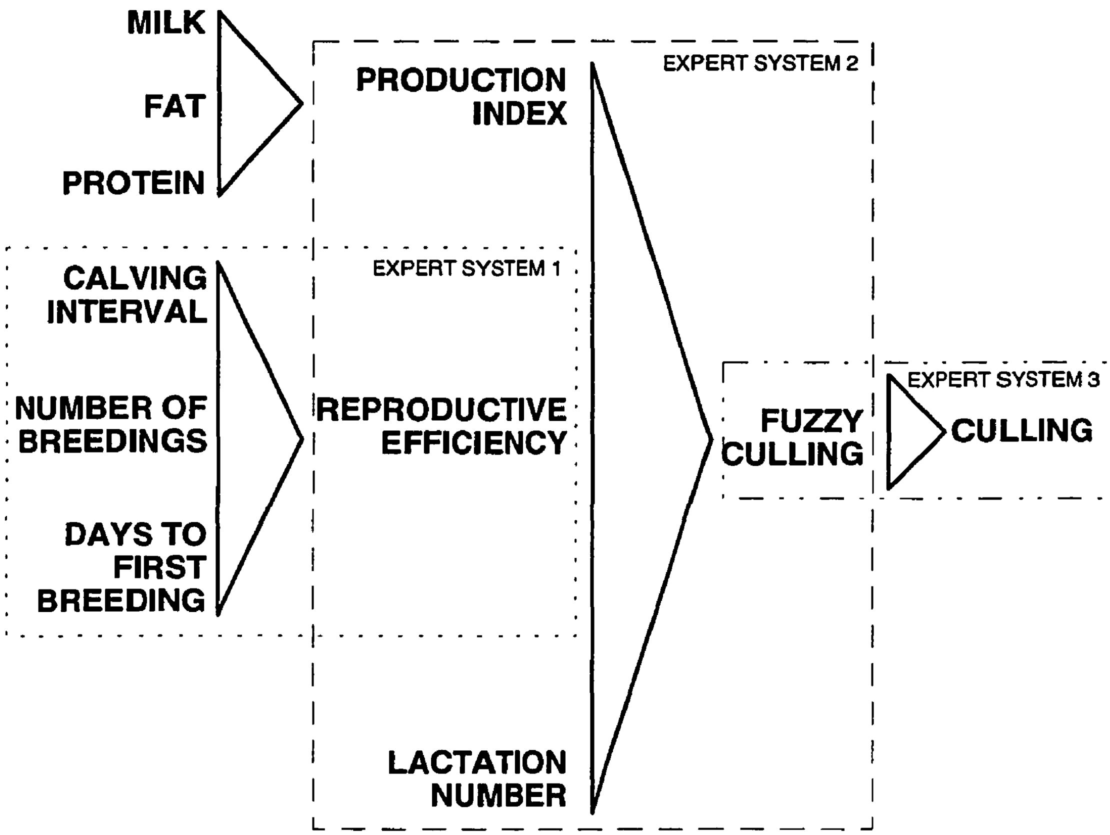

Figure 2.3 - Decision-tree of variables and expert systems in the prototype decision-support system.

MILK PRODUCTION FUZZY SETS

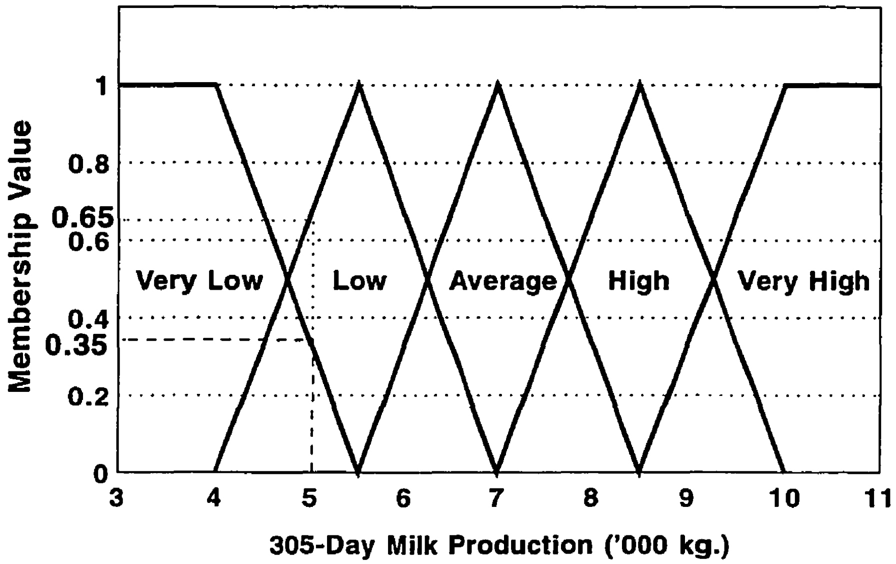

Figure 2.4 - An hypothetical group of fuzzy sets representing the variable "305-day milk production".

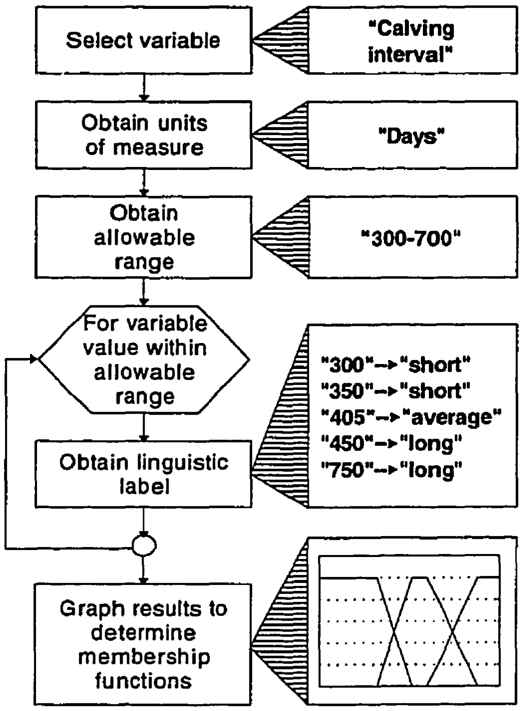

Figure 2.5 - An illustration of the first procedure for the elicitation of membership functions.

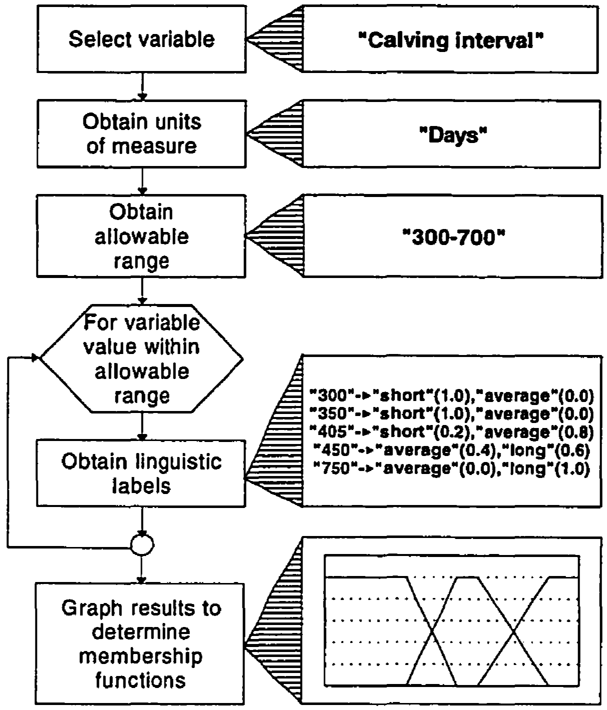

Figure 2.6 - An illustration of the second procedure for the elicitation of membership functions.

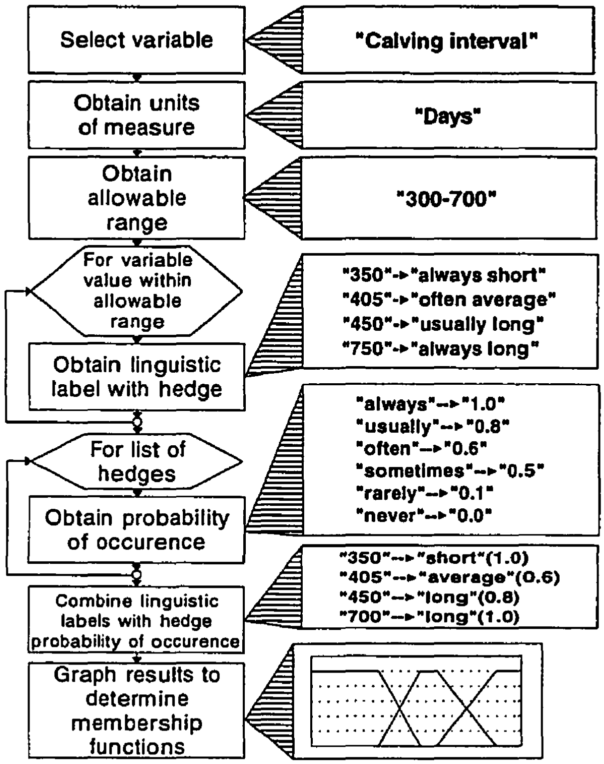

Figure 2.7 - An illustration of the third procedure for the elicitation of membership functions.

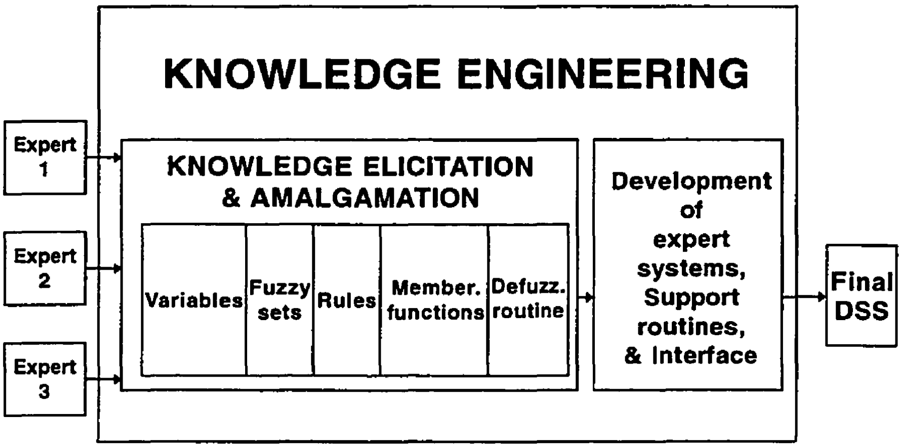

Figure 2.8 - The amalgamated decision-support system construction process.

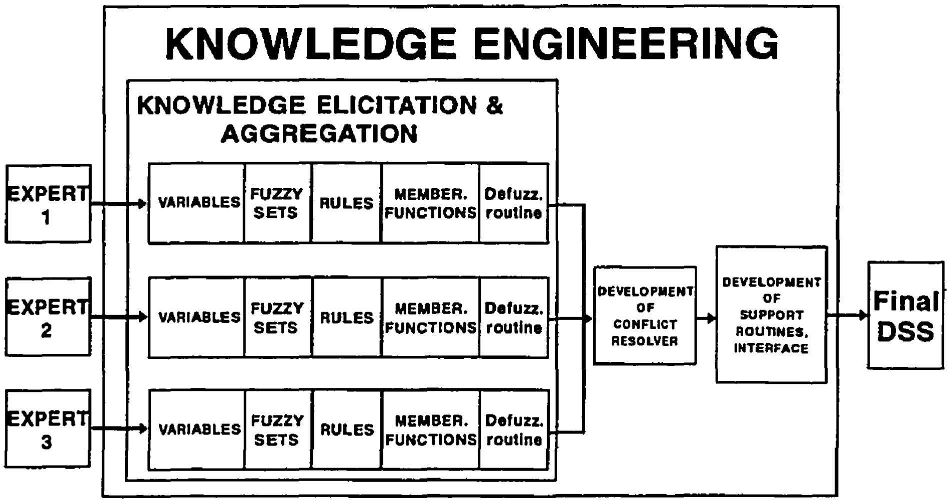

Figure 2.9 - The aggregated decision-support system construction process.

CONNECTING STATEMENT

As a first step towards the development of a global decision support system for dairy cattle breeding management, a first-generation DSS was constructed with the objective of recommending culling decisions. In this prototype, the knowledge of three experts was amalgamated into a single system and fuzzy logic was used to complement the traditional decision-support expert system approach. The process of prototyping served a number of purposes: firstly, it allowed us to gain some insight into the complexity of the area of interest or the "domain". This first-generation prototype was limited in scope, considering a total of five input variables. Upon further analysis it was determined that the amalgamation approach to DSS development had the shortcoming of potentially not representing the knowledge of any particular expert. For these reasons, it was determined that the creation of a second-generation prototype should be pursued following an aggregated approach. The aggregation approach is structured in a format so that the knowledge of each expert is maintained independently and is encoded in individual modules.

Chapter 3

A Decision Support System for the Recommendation of Dairy Cattle Culling Decisions

ABSTRACT

One of the primary objectives of all dairy production enterprises is to increase income while reducing expenses. This can be partly accomplished through improved management practices. Since it is possible to automate optimal decision-making through the use of decision-support systems, expert knowledge becomes available to all. One area which would benefit from the "packaging" of expert knowledge is the dairy cattle breeding sector. Dairy cattle breeding requires continuous decision-making in many areas such as culling, replacement, and mating. The objective of this study was to develop a second-generation prototype decision-support system for culling in dairy cattle. Contrary to previous work where the encoded expertise was limited to three experts from the same domain, this research used two individuals from each of three domains (Dairy researchers, Producers, and Dairy herd improvement specialists). An aggregation approach was used which involved the development and maintenance of separate modules, each containing the compiled expertise of one of the six experts. Although all experts had a common overall goal of culling, it was observed that their intermediate objectives and methods in achieving the overall goal varied considerably. For example, two experts focused primarily on low milk production as a factor for culling while the other four also considered conformation information before making a decision. Although six functional modules have already been developed and are capable of recommending culling decisions independently, it was envisaged that these modules would be combined in order to develop a single aggregated decision support system. This is due, in part, to a number of factors including significant "overlap" between the considerations of the experts, requiring a greater level of system maintenance, differing conclusions which make aggregating modules difficult, difficulty in validating such a system since the intermediate objectives of the experts differ and, therefore, the culling decisions are not expected to be uniform either.

(**Key words**: decision-support system, culling, dairy)

INTRODUCTION

Profitability, the main objective of a dairy farming enterprise, is directly related to the everyday expenses and income of the farm operation (Mainland, 1994). Almost all dairy farmers try to increase their income by increasing milk yield, which also serves the purpose of diluting maintenance and other fixed costs (McDaniel, 1994). This is often accomplished through the removal of lower producing cows and replacing them with higher producing ones. Economically, it has been determined that the "top 20%" of producers have significantly lower replacement costs compared to the "bottom 20%" (Pellerin, 1993). Therefore, if it is possible to acquire expertise from these "top" producers, or other experts, and make this expertise available for use on less productive farms. This would potentially improve the profitability of the latter, which would also complement the knowledge of other experts, for example, advisors. Unfortunately, human experts are costly and in short supply (Levins and Varner, 1987) and it is also difficult to transfer or impart their expertise. Furthermore, budget cuts adversely affect the availability and quality of experts (Willett and Andrews, 1996). It is therefore suggested that these last problems can be alleviated with the development of computer systems which could be used to mimic the reasoning processes of experts (Willett and Andrews, 1996; Levins and Varner, 1987). Decision-support systems (DSS) have proven to be effective tools which can be used to capture and convey expert knowledge (Cox, 1996; Allore *et al.*, 1995).

One area which would benefit from the "packaging" of expert knowledge is the dairy cattle breeding sector. Dairy cattle breeding requires continuous decision-making in many areas such as culling, replacement, and mating. As a first step towards the development of a global decision support system for dairy cattle breeding management, a first-generation DSS was constructed with the objective of recommending culling decisions (Lacroix *et al.*, 1997). In this prototype, the knowledge of three experts was amalgamated into a single system and fuzzy logic was used to complement the traditional decision-support expert system approach. The process of prototyping, which is often the first stage in DSS development (Smith, 1989), served a number of purposes: firstly, it allowed us to gain some insight into the complexity of the area of interest or the "domain". Also, since DSS are domain-specific there are no absolute formulae for developing such systems; the prototype was analyzed so that a more thorough, or more domain-specific, methodology for DSS development was established (Strasser *et al.*, 1997). This first-generation prototype was limited in scope, considering a total of five input variables. Upon further analysis it was determined that the amalgamation approach to DSS development had the shortcoming of potentially not representing the knowledge of any particular expert (Strasser *et al.*, 1997). For these reasons, it was determined that the creation of a second-generation prototype should be pursued. The objective of this study was to develop a second generation culling decision-support system prototype. Specifically, the objectives were:

- to acquire expert knowledge using the procedures outlined in the framework (Strasser *et al.*, 1997), expanding the expertise to include the knowledge of six experts from three different domain areas,
- to develop individual DSS modules for each expert, whereby each expert's knowledge would be maintained independently from the knowledge of other experts,
- to examine the possibility of integrating the six modules into a single aggregated DSS,
- and to pursue the evaluation of the applicability of fuzzy logic in DSS.

EXPERIMENTAL PROCEDURE

The process of developing a DSS can be characterized in four stages: the groundwork stage, knowledge elicitation, knowledge implementation, and system evaluation (Strasser *et al.*, 1997). The groundwork stage includes the decisions which must be made before knowledge elicitation can commence. Knowledge elicitation is the process of knowledge transfer from the expert to the knowledge engineer. Knowledge implementation is the process of interpreting and encoding the knowledge obtained. Lastly, system evaluation serves the purpose of measuring the accuracy and relevancy of the system developed.

Project Groundwork

Since the goal of the study was predetermined (i.e., to create a decision-support system capable of recommending culling decisions), the first objective was the selection of appropriate domain experts. Two domain experts were selected from each of three expert-types: dairy producers, researchers, and DHI specialists. The first expert-type consisted of two dairy producers chosen from a management index listing of the "most profitable" producers in Québec as supplied by the local dairy herd improvement agency (DHI). These experts were selected for their applied level of expertise. The second expert-type, two researchers, were chosen because of their continuing research in the area of dairy cattle breeding and, in particular, for their theoretical knowledge applied to dairy cattle culling. Finally, two DHI specialists were selected. It was hoped that the DHI specialists would offer an intermediate level of knowledge--a level of knowledge which would serve as a link between the knowledge levels of the two other expert-types (Figure 3.1).

Knowledge Elicitation

Knowledge elicitation sessions consisted of interviews with each of the six experts. Prior to the first interview each expert was contacted and the objectives of the interview were outlined. Each interview session consisted of a period of informal discussion and a period of formal interview. The objective of the period of informal discussion was to introduce the expert to the project, in the case of the first interview, and re-familiarize the experts with the project, as was the case during subsequent interviews. Informal discussions were of a duration of 0.5 to one hour. The formal interview consisted of two sessions conducted with four of the experts (the producers and researchers) with each session requiring one to 2.5 hours. Only one interview was conducted with the other two experts (the DHI specialists) combining the procedures used in interviewing the other experts into one session. In one case the interview session lasted 3.5 hours and in the other the session required 1.5 hours. In total, approximately 18.5 hours were spent in formal interviewing.

The objectives of the first interview were to formally introduce the project to the experts and to obtain the variables that they consider in making a culling decision. This first interview followed an unstructured interview format; the experts were permitted to follow their own ideas with only minimal guidance from the knowledge engineer. The follow-up interview was more structured. The experts were presented with their list of the variables, obtained from the audio recordings of the first interview, and were asked to confirm their validity. This was followed by direct questioning in order to obtain the relationships between the variables. Some of the variables considered were "fuzzy"; in other words the input value for the variable was quantitative but the terms used to describe the variable were qualitative. For example, Expert 1 described milk production with qualifier terms such as "good", "ok", and "bad", instead of an actual quantitative value for milk production. Since one of the objectives was to develop individual modules using the terminology of the individual experts, which is also known as an expert's "personal language" (Strasser *et al.*, 1997), it was necessary to use a technique for "translating" the quantitative input values to the qualifiers used within the respective modules. In fuzzy set theory, this process of translation is called *fuzzification* (Lacroix *et al.*, 1997).

Fuzzy set theory provides a mathematical framework for the manipulation of vague and imprecise information. With fuzzy mathematics, sets are used to represent concepts, and membership values are used to denote to what degree a value can be associated to a particular fuzzy set. Fuzzy sets correspond to qualifiers (which are also known as linguistic labels). This degree of association is represented by a membership value (MV) to the fuzzy set. Membership values often range from zero, denoting no membership to a fuzzy set, to 100, representing complete membership to a fuzzy set. For simplicity, in this research, fuzzy sets were represented by triangles and trapezoids with the summits of the triangles and the points of discontinuity in the top segments of the trapezoids represented by "critical points" (see Figure 3.4). These critical points were obtained through direct questioning of the expert. A more complete description of fuzzy logic and the mathematics of implementing fuzzy logic within expert systems can be found in Lacroix *et al.*(1997).

Knowledge Implementation

Knowledge implementation refers to the process of encoding in a knowledge base, the knowledge of experts obtained through the knowledge elicitation process. Since this project considers the knowledge of multiple experts with different intermediate objectives, it was determined that their respective knowledge should be aggregated. This ensures that the knowledge of each expert is kept separate as opposed to amalgamation, whereby the knowledge of multiple experts is combined to create a single knowledge-base (Figure 3.2) (Strasser *et al.*, 1997). The aggregation approach is structured in a format so that the knowledge of each expert is maintained independently and is encoded in individual modules (Figure 3.3). Since one of the objectives of this research was to apply the aggregation approach towards the construction of a DSS, the modules were developed using the variables and the relationships between the variables considered by the individual experts. When the knowledge was expressed heuristically, for example, "If the milk production is low then cull the cow", the knowledge was implemented in the form of production rules, which is the most commonly used method for knowledge implementation (McKinion and Lemmon, 1985). Knowledge was also expressed by the experts in a procedural fashion, for example, "First, I calculated the economic value of milk, then I added protein and fat, to obtain a score for the cow related to production". Under these circumstances the knowledge was implemented in a procedural format.

System Evaluation

In order to ensure that the goal of a DSS has been reached, evaluation must be performed. An expert system evaluation consists of measuring the system's accuracy and usefulness (Batchelor and McClendon, 1989). The two phases of evaluation considered for this prototype consisted of a process of verification and an analysis of the modules. Verification was conducted to ensure that the expert system performs as intended, by searching for missing rules, dead-end IF conditions, mistyped rules, etc. This was performed by operating the modules using the data of 33 dairy cows, obtained from the local university animal research facility. If the module outputs did not appear reasonable, the results were traced backward from the decision to determine the problem. The analysis of the modules consisted of an examination of the knowledge and the approach used by each expert. In doing so, the variables, the rules, and the dairy management areas considered by the experts were examined. An analysis was also performed to determine the appropriateness of employing fuzzy logic to represent the experts' reasoning. Finally, the possibility of integrating the individual modules into a single aggregated DSS was also evaluated. This was explored focussing on the compatibility, contradictions, and the overlap between the modules. This resulted in an overall appraisal of the applicability of the aggregation approach.

RESULTS

Module Descriptions

Since one of the overall project objectives was to follow an aggregated approach to DSS development, a total of six modules were constructed; one for each of the six experts consulted. As has been previously stated, all the experts were allotted the same objective, i.e., to perform culling decisions. Each expert was permitted to consider any variable which they felt was relevant in order to arrive at a culling decision. The types of variables considered can be summarized in four categories: qualitative, fuzzy, quantitative, and logical. Qualitative variables are those variables which are described by a quality. For example, conformation traits are often described with such terms as "very good", "good", "poor", etc. Qualitative traits are not based on quantitative values and often represent the subjective assessment of the variable. Fuzzy variables are those variables which are described with a quality but are based on an actual numerical value. One example is milk production which can be described by terms such as "low", "high", etc., but which refers to a numerical value (e.g., kilograms of milk). Quantitative variables are described by their numerical value, for example, lactation number. Finally, logical variables are those variables which can have a value of "yes" or "no".

The following sections contain a description of each module which was developed. In particular, the culling decision is examined followed by a description of the variables. Also described is the general structure of the knowledge embedded within each module. Then an overall evaluation of the modules is performed.

Module One

The objective of Module One, created from the knowledge of a producer expert, is a "Yes" or

"No" culling decision. This module is divided into six sections which represent various cow level management areas: SCC & Mastitis, Udder, Feet & Legs, Production, Reproduction, and Miscellaneous (Figure 3.5). Each section performs an evaluation of that particular management area. The results obtained from each section are introduced into a set of rules which recommends a culling decision. To arrive at a culling decision a total of 22 variables were consulted (Table 3.1). Of these, 8 were described as qualitative, 3 were fuzzy, 2 were quantitative, and 9 were logical. Most of the qualitative and logical variables are obtained through a visual appraisal of the cow that is to be evaluated. As such, these variables are not available from traditional databases, thus requiring the user to assess the animal and enter the value in the system. Therefore, the expertise of the user in making such an evaluation will influence the quality of the decision. As previously mentioned, the three fuzzy variables must be fuzzified for use within the module. Three fuzzy sets were considered sufficient to characterize each of the fuzzy variables. The fuzzy sets for the production variables were described with the qualifiers "Bad", "Acceptable", and "Good". The critical points were chosen at 7000, 7500, and 8000 kg. 305-day milk production for first-lactation heifers (Figure 3.4) and 7500, 8000, and 9000 kg. 305-day milk production for cows in their second lactation or greater. The qualifiers used to name the fuzzy sets for Somatic Cell Count (SCC) were "Very Good", "Good", and "Bad" with the critical values chosen at 250000, 350000, and 500000 units.

Each section of Module One is constructed to obtain the variables and to produce a decision relative to that particular area. For example, each section consists of routines for obtaining the relevant input data, fuzzifying this data (as with the SCC & Mastitis and Production sections), consulting a set of rules to obtain a value for the section (either "Keep" or "Cull"), and making accessible this value to allow for an overall system culling decision. A total of six sets of rules were constructed; one for each section, except the Miscellaneous section, where no rules were considered, and one overall set of rules recommending a culling decision.

Module Two

Module Two, developed from the expertise of the second producer expert, will not only recommend the expected "Yes" or "No" culling decision, but other recommendations are also expressed. These recommendations include "Sell", "Watch", Surrogate", and "Unknown". "Sell" referred to a decision to remove the animal from the herd with the intention of selling it to another producer. This is only considered if the animal has poor production relative to the herd, but has no health problems and the conformation is very good. "Watch" is a flag condition recommended when the cow has shown signs of physical problems which might deteriorate over the next lactation. "Surrogate" is recommended if the cow has good production but is lacking in conformation. She can be used as a potential embryo recipient. An "Unknown" decision will result if the culling decision cannot be made with a high degree of certainty. It is obvious from these "extra" recommendations that the expert consulted for the development of this module was actively involved in the production and sale of animal genetics. The implications of this will be discussed later.

The module was divided into six management areas: mastitis, udder, feet & legs, production, reproduction, and miscellaneous (Figure 3.6). As with Module One rule sets are consulted so that an intermediate decision is made representing each management area. An overall rule set recommends the culling decision.

Module Two considered 17 input variables, consisting of 3 qualitative variables, 3 fuzzy variable, 3 quantitative variables, and 8 logical variables (Table 3.2). Upon examination of the variables it is evident that all the conformation traits considered were expressed in the form of qualitative variables. The logical variables were mostly considered in order to obtain an overall understanding of the health of the animal. The fuzzy variables, temperament and milking speed, are fuzzified not through critical points but rather, the variables were assigned to particular fuzzy sets. For example, a score of 1 was fuzzified as "bad" with a membership value (MV) of 100, and 2 as "bad" with an MV of 66, etc. The fuzzification of the production variable, breed class average for milk, was accomplished through the use of membership functions with the critical points defined at the herd breed class average (BCA), for the triangular fuzzy set, and herd BCA ± 45 herd BCA for the other two critical points.

Module Three

As with the first module, the objective of Module Three, obtained from the knowledge of a DHI specialist, is to recommend a "yes" or "no" culling decision. Module Three is divided into five sections; SCC & Mastitis, Udder, Feet & Legs, Production, and Reproduction (Figure 3.7). Rule sets are consulted to evaluate SCC & Mastitis, and Production. The udder is evaluated by analysing the characteristics of the udder. A cow having a major udder problem, for example, a blind quarter, will obtain a value of "cull" for the udder management area. Major feet & leg problems, such as chronically swollen hocks, will result in a decision to cull for feet & legs. A decision to cull the cow for reproduction is based primarily on whether the cow has surpassed 200 days open or has gone dry without being successfully re-bred. An overall decision to cull will result if a culling recommendation is suggested for one of the management areas.

This module consists of 28 input variables (Table 3.3). Two variables were described as qualitative, 5 as quantitative, and 21 as logical. No fuzzy variables were considered. Logical variables were used to determine the health status of the animal and to evaluate the cow's conformation. Quantitative variables were related primarily to production and SCC & mastitis.

Module Four

Module Four was developed from the expertise of the second DHI specialist. Unlike the other modules developed, Module Four does not recommend a decision for culling. This module produces a ranking for each cow requiring the user to make the final decision. This module requires the user to enter the variables for each cow. Module Four is organized into four management areas: Health, Conformation, Production, and Genetics (Figure 3.8). The culling decision is derived from a mathematical procedure rather than from rules. The input variables linearly combined (30% Health, 30% Production, 25% Conformation, and 15% Genetics), resulting in an overall performance score for each cow.

A total of four quantitative input variables were considered (Table 3.4). Three of the four variables, “health”, “conformation”, and “genetics” are acquired by requesting that the user enter a value between 0 and 100 representing the cow's performance relative to the herd average. A value greater than 50 would indicate that the cow exceeds the herd average for that particular variable. A value less than 50 indicates a level of performance below the herd average. The fourth variable, “cow rating” is obtained from the local DHI records. Cow rating is based upon 4% fat corrected milk adjusted to a mature equivalent basis established on a within-herd basis. It should be noted that this module was devised on the assumption that many producers have a good understanding of how a particular cow compares to the herd without having to resort to records.

Module Five

The objectives of Module Five is to provide a listing of cows which should be culled, if a culling rate is specified by the user, or to rank the cows when no culling rate is specified. The module is divided into four management areas: health, production, and miscellaneous considerations (Figure 3.9). A decision to cull results if a cow has a major problem with health. If there are no major health problems and a cow is in the third lactation at least, she is ranked according to her production records.

This module was developed from the knowledge of a researcher and considers a total of eight input variables (Table 3.5): five quantitative variables and three logical variables. There are no fuzzy or qualitative variables considered. Quantitative variables are used to evaluate production and to determine the number of cows to cull if a culling rate is specified by the user. The logical variables are used to determine the overall health of the cow.

Module Six

The sixth module, encoded from the knowledge of the second researcher, is another procedural module, like Module Four, with the objective of ranking the cows. The decision to cull is left to the discretion of the user. The module is divided into three management areas: production, reproduction, and miscellaneous costs (Figure 3.10). The intermediate objective of this module is to create an index which would roughly simulate the average net income of a cow. A final rating is calculated by linearly combining production, reproduction, and miscellaneous costs using the weights 0.55, 0.35, and 0.10 respectively, as suggested by the expert, which represent the relative economic importance of each management area. The expert expressly stated, however, that actual economic values do not currently exist at this time, but when they become available they should be included within the module.

A total of eight quantitative variables were considered (Table 3.6). In order to obtain a unitless variable from each management area all the input variables are divided by the respective herd standard deviation.

Overall Results

An examination of the results indicates that the experts differed in their culling objectives. Whereas all the experts had the goal of recommending culling decisions, this was not limited to a "keep" or "cull" ("yes" or "no") decision. A consultation with two of the six modules, modules Four and Six, results in a ranking of the cows examined, requiring the user to make the final decision. Furthermore, one module, Module Two, introduces other possibilities such as "Sell", "Watch", or "Surrogate". This demonstrates that although all the experts had a common goal their intermediate objectives differed. This is further demonstrated through an evaluation of the variables considered by the various experts.

There was a variety of opinion as to which variables should be considered for a culling decision. One common thread between all the modules is the consideration of production, although there were differences in the variables used to evaluate production. Three experts (i.e., the two DHI specialists and Researcher 1) considered the variable “cow rating” to evaluate an animal’s production. Producer 1 was only interested in 305-day milk production. Producer 2 considered the cow’s breed class average for milk and the second researcher considered the 305-day production values for milk, fat, and protein. Another common thread between the modules was the consideration of a cow’s health status. Most of the modules expressly requested data related to health, (e.g., the consideration of feet & leg problems). Module Six was an exception, given that the sixth expert felt that the health status of the cow was implicitly expressed by the other variables considered. For example, a longer calving interval may be an indication that the cow is having reproduction difficulties. Other differences existed in the considerations of the experts. For example, the producers and the DHI specialists all considered conformation traits in making a culling decision, either through explicit conformation variables, (e.g., "udder conformation score" considered in Module Two), or through implicit conformation variables, (e.g., "good fore attachment" in Module Three). The researchers did not consider conformation at all.

As previously stated, the various input variables were categorized into four categories: qualitative, fuzzy, quantitative, and logical. Many qualitative and logical variables, for which data is not maintained in traditional databases, were considered in the reasoning processes of the experts. For example, Producer Expert 1 believed that a cow which has trouble laying down or getting up is an indication that the cow has a deteriorating feet & leg condition, which may not be evident through a visual inspection of the animal. Of course, this consideration is not maintained as a variable by any external source (e.g., DHI agencies). Most of the quantitative variables were related to data which were readily available through DHI agencies (e.g., milk production or somatic cell count). Fuzzy variables were considered less frequently than was anticipated at the outset of the project. The only two modules which considered fuzzy variables were those developed from the expertise of producers.

Through an overall examination of the modules developed it can be seen that all experts arrived at an overall decision through an evaluation of individual management areas. All experts gave consideration to the management areas of reproduction, health, and production. Conformation was only considered by the DHI specialists and the producers but not the researchers. The researchers suggested that the primary objective in culling is to improve production; consideration given to conformation would be more appropriate when determining a suitable replacement for the cow culled.

From the results it was also evident that there were similarities in the considerations of experts of the same expert-type. For example, the producers examined the same management areas and placed a significant level of importance upon the visual appraisal of their animals. This is apparent from their inclusion of conformation traits. The DHI specialists showed similarities in the techniques used in evaluating animals. Since they do not have first-hand experience with particular animals they obtain much of the necessary information through a direct questioning of the farm manager, the exception being production data which they obtain through their DHI agency. This factor is represented within the modules through the inclusion of many variables which require input from the user. Indirectly, both modules elicited from the DHI specialists require the user to evaluate their animals. Finally, the researchers both demonstrate the prioritization of production variables in their evaluation of the cows.

DISCUSSION

Based upon the experience obtained through the development of the first-generation prototype one of our objectives was to include, within this second-generation prototype, the knowledge of multiple expert-types, in this case, researchers, DHI specialists, and dairy producers. Since it was expected, and desired, that each expert would contribute their own priorities to the project, it was suggested that the acquired knowledge be implemented using an aggregated approach. Aggregation requires the individual development of modules representing the knowledge of each expert. Through the creation of these independent modules, the experts' intermediate objectives can be maintained and made available to the user. Although the aggregation approach to DSS development has many advantages over the amalgamation approach, (e.g., the DSS is easier to update as new knowledge becomes available), the results obtained from this project have demonstrated that there also exist some limitations to this approach. For example, from a development perspective, a greater number of variables need to be managed. Also, some of these variables may not be obtainable through traditional databases, (e.g., whether the cow has trouble laying down), and would, therefore require user input. Another consideration is that many variables are unavailable to many producers. Those producers who do not classify their animals will not have any conformation records available.

Also, producers who are not registered with the local DHI will not have access to "treated" production variables such as, cow rating, or breed class average. Another consideration regarding variables is that if there are too many variables to input, the system developed will not be as readily accepted by the targeted user group. Also, since modules are constructed for each expert, with the required development time associated with the construction of each, this can detract from the perfection of a smaller system.

As was seen from the results, the intermediate objectives of the experts differed. This was demonstrated through the different possible conclusions the experts could arrive at in making a culling decision and in the variation in the variables considered. This also resulted in potentially conflicting evaluations of each particular management area. One of the criteria for aggregation is that the overall objectives of each module be the same. As was demonstrated this was not the case. As it exists the experts were permitted to consider any management area and therefore, considerable "overlap" was evident. Aggregation might be best implemented if most of the overlap between the knowledge of the experts is minimized. For example, DHI specialists would be limited to making decisions related to production, veterinarians would be limited to decisions related to animal health and welfare, and breed agency specialists would be limited to conformation. This might "improve" the overall quality of the knowledge acquired, limit the number of variables in the DSS, and prevent conflicting evaluations. From this, overlap would still exist but it would be minimized.

There exists another reason why it would be extremely difficult to aggregate the existing modules into a single DSS. Since each expert had their own intermediate objectives, it would be very difficult to evaluate such a DSS in great detail. Since each expert brings his/her own priorities to the project, the overall decisions will not coincide. Therefore, an overall system validation cannot focus on the comparison of the decisions between experts. Other procedures become necessary. A more appropriate procedure for the validation of DSS would be field testing, which is conducted by the potential users of the system. This helps establish the acceptable performance ranges. It also allows for the immediate revelation of any weaknesses in the flow of communication between the system and the user (Harrison, 1991). This procedure will be employed in a subsequent stage of this research project.

Concerning the selection of the experts, it was evident from the results that the producers demonstrated an interest in the consideration of conformation traits. This was not surprising since these producers generated a portion of their income through the sale of animals. For this reason the management practices of these producers may not reflect the management practices of a producer engaged solely in the production of milk. Unfortunately, this factor was not considered at the outset of the project and only became apparent after the project was underway. It was then hoped that since the DHI specialists have accumulated experience in providing expertise to both those involved in the sale of animals and those involved primarily in the sale of milk, their inclusion would allow for the addressing of the priorities of the latter group. Any future DSS must more closely consider the background of the experts to avoid any conflicts of interest.

Another policy consideration is related to the politics and economics of DSS development. As the current economic situation in Québec evolves, there is much talk related to the phasing-out of the system of subsidies as is required by the various economic trade agreements which influence the dairy sector, including the Global Agreement on Tariffs and Trade (GATT), and the North American Free Trade Agreement (NAFTA). As the economic climate changes it is expected that this will have an influence on culling strategies. For example, under the Canadian quota system, it is common to see that once the maximum permitted production levels are attained many producers place a great deal of importance on the improvement of conformation traits. As the production ceiling is eliminated it is certain that this will have a net effect on the strategies associated with culling. The goals of future DSS will be affected by any changes in policy. Therefore, future DSS must be constructed to be adaptable to the changes which evolve.

It is obvious that experts play a primary role in DSS development. It is therefore important to analyse their contributions. At the expert level, there are two main reasons which hindered the gathering of knowledge. Firstly, the six experts were not remunerated. Although they were more than willing to donate their time to the project, their primary objectives were their respective concerns for their individual work. As such, there was always some difficulty in scheduling appointments and there was always an underlying concern, and reasonably so, that the number of interviews necessary should be minimized so as not to interfere with their personal work. Also, for some of the experts who were unfamiliar with the research environment, the level of questioning associated with the knowledge acquisition interviews may have been a little intimidating. It was, therefore, not always possible to adhere exactly to our developmental framework. This was expected since the creation of a global framework, applicable in all circumstances, is a formidable task.

One of the overall objectives of this research was to determine the applicability of applying fuzzy logic to DSS development. The results demonstrate that the use of fuzzy variables were limited to two experts; the producers. This may be because the producers are so accustomed to their work that they do not actively consider precise information. Therefore, they would more likely express certain concepts, such as production, in fuzzy terms. This is not the case with other experts who often require more precise information before arriving at a decision. For example, as with the researchers and the DHI specialists, an evaluation of production required actual quantitative values for the cow as well as quantitative values for the herd as a comparison. These variables were treated as quantitative throughout the decision-making process. The consideration of fuzzy logic might be most appropriate when implementing knowledge from a source which is so close to the domain that they do not require explicit precision from the variables that they consider.

CONCLUSIONS

Six modules were developed with the objective of recommending culling decisions. Although six functional modules have already been developed and are capable of recommending culling decisions independently, it was envisaged that these modules would be combined in order to develop a single aggregated decision support system. In fact, this has proven difficult, due, in part, to a number of factors:

- There exists a significant amount of "overlap" between the considerations of the experts, requiring the maintenance of a significant amount of variables.
- Since the conclusions of each module are not equivalent, it would be difficult to combine modules.
- Validation of this DSS would be extremely difficult because the intermediate objectives of the experts are not the same and, therefore, the culling decisions are not expected to be uniform either.

For these reasons, future decision support systems should limit the acquisition of expert knowledge to fewer experts, or limit the amount of "overlap" permitted between experts. The aggregated approach has many advantages but requires greater uniformity between modules developed. Finally, the applicability of integrating fuzzy logic with decision support system development is limited. Many of the variables considered by the experts are unavailable through traditional databases, and many other variables are already described qualitatively. Therefore there is no need for the *translation* which is accomplished through the use of fuzzy logic.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada, Semex Canada and the Cattle Breeding Research Council of Canada. Data for evaluation were provided by the Macdonald Campus farm.

REFERENCES

Allore, H.G., L.R. Jones, W.G. Merrill, and P.A. Oltenacu. 1995. A decision support system for evaluating mastitis information. Journal of Dairy Science. 78:1382-1398.

Batchelor, W.D. and R.W. McClendon. 1989. Two approaches to knowledge engineering in an insect pest management expert system. American Society of Agricultural Engineers.

Paper 89-7081.

Cox, P.G. 1996. Some issues in the design of agricultural decision support systems. Agricultural Systems. 52:355-381.

Harrison, S.R. 1991. Validation of agricultural expert systems. Agricultural Systems. 35:265-285.

Lacroix, R., M. Strasser, R. Kok, and K.M. Wade. 1997. Decision support systems for culling of dairy cows using fuzzy logic. Submitted to the Journal of Dairy Science. March 1997.

Levins, R.A., and M.A. Varner. 1987. An expert diagnostic aid for reproductive problems in dairy cattle. Computers and Electronics in Agriculture. 2:47-56.

Mainland, D.D. 1994. A decision support system for dairy farmers and advisors. Agricultural Systems. 45:217-231.

McDaniel, B.T. 1994.Genetic evaluations of secondary traits in dairy cattle. Proceedings of the 5th World Congress on Genetics Applied to Livestock Production. 17:59-60.

McKinion, J.M. and H.E. Lemmon. 1985. Expert systems for agriculture. Computers and Electronics in Agriculture. 1:31-40.

Pellerin, D. 1993. L'Elevage des genisses: La moderation a bien meilleur cout. Symposium sur les Bovins Laitiers. 28 Octobre 1993. Conseil des Productions Animales du Québec. p. 13-57.

Smith, T.R. 1989. The potential application of expert systems in dairy extension education. Journal of Dairy Science. 72:2760-2766.

Strasser, M., R. Lacroix, R. Kok, and K.M. Wade. 1997. A framework for the development of fuzzy decision support systems. Submitted to the Journal of the Canadian Society of Agricultural Engineering. March 1997.

Willett, M.J. and P.K. Andrews. 1996. Knowledge-based systems for fruit and vegetable production management. AI Applications. 10:75-92.

MANAGEMENT AREA	VARIABLE	TYPE	ALLOWABLE VALUES
SCC & Mastitis	SCC	Fuzzy	1-9000000
	Stage of Lactation	Qualitative	early, middle, late
	No. of Mastitis Treatments	Quantitative	0-10
Udder	Four Functioning Teats	Logical	yes/no
	Fore Udder	Qualitative	long/bulgy
	Pendulous Udder	Logical	yes/no
	Rear Udder	Qualitative	wide/narrow
	Teat Length	Qualitative	normal/abnormally long
	Good Median Suspension	Logical	yes/no
	Good Teat Placement	Logical	yes/no
Feet & Legs	Rear Leg Shape	Qualitative	good/bad
	Heel Angle	Qualitative	straight/crooked
	Hock Size	Qualitative	small/bulgy
	Walking Characteristics	Qualitative	good/bad
	Trouble Getting Up or Laying Down	Logical	yes/no
Production	Quota Level Attained	Logical	yes/no
	Lactation Number	Quantitative	1-20
	305-day Milk Production (Most Recent)	Fuzzy	1-30000
	305-day Milk Production (Projected)	Fuzzy	1-30000
Reproduction	Number of Inseminations > Herd Average	Logical	yes/no
	Solid Back	Logical	yes/no
Miscellaneous	Miscellaneous Factors Influencing Culling	Logical	yes/no

Table 3.1 - Variables considered in Module One (obtained from Producer 1).

MANAGEMENT AREA	VARIABLE	TYPE	ALLOWABLE VALUES
Mastitis	Mastitis (Current)	Logical	yes/no
	Mastitis Curable	Logical	yes/no
	Recurring Problems with Mastitis	Logical	yes/no
Udder	Udder Conformation Score	Qualitative	excellent, very good, good plus, good, fair, poor
	Udder Problems	Logical	yes/no
Feet & Legs	Feet & Leg Conformation Score	Qualitative	excellent, very good, good plus, good, fair, poor
	Current Non-Treatable Feet & Leg Problems	Logical	yes/no
Production	Herd Breed Class Average	Quantitative	1-999
	Cow Breed Class Average	Fuzzy	1-999
	Physical Problems Affecting Production	Logical	yes/no
Reproduction	Number of Inseminations	Quantitative	1-20
	Final Class (Conformation)	Qualitative	excellent, very good, good plus, good, fair, poor
Miscellaneous	Milking Speed	Fuzzy	1-5
	Temperament	Fuzzy	1-5
	Special Considerations	Logical	yes/no
	Lactation Number	Quantitative	1-15
	Embryo Donor	Logical	yes/no

Table 3.2 - Variables considered in Module Two (obtained from Producer 2).

MANAGEMENT AREA	VARIABLE	TYPE	ALLOWABLE VALUES
SCC & Mastitis	Number of SCC Tests (Current Lactation)	Quantitative	0-∞
	First Test Within One Week of Calving	Logical	yes/no
	High SCC Due to Acute Mastitis	Logical	yes/no
	Mastitis Cured	Logical	yes/n
	Mastitis Incidences (Current Lactation)	Quantitative	0-12
	Mastitis Always Cured	Logical	yes/no
Udder	Solid Udder	Logical	yes/no
	Udder Too Voluminous	Logical	yes/no
	Good Fore Attachment	Logical	yes/no
	Pendulous Udder	Logical	yes/no
	Recurring Edema	Logical	yes/no
	Blind Quarters	Logical	yes/no
	Equal Sized Quarters	Logical	yes/no
	Good Rear Attachment	Logical	yes/no
	Strong Median Suspension	Logical	yes/no
	Good Teat Placement	Logical	yes/no
	Teat Defects	Logical	yes/no
	Teat Damage	Logical	yes/no
Feet & Legs	Rear Leg Curvature	Qualitative	good/bad
	Heel Angle	Qualitative	good/bad
	Hock Problems	Logical	yes/no
Production	Cow Rating	Quantitative	60-140
	Lactation Number	Quantitative	1-20
	Calving Difficulties Affected Production	Logical	yes/no
Reproduction	Days Open	Quantitative	1-600
	Cow Dry	Logical	yes/no
	Embryo Donor	Logical	yes/no
	Very Good Overall Conformation	Logical	yes/no

Table 3.3 - Variables considered in Module Three (obtained from DHI Specialist 1).

MANAGEMENT AREA	VARIABLE	TYPE	ALLOWABLE VALUES
Health	Health	Quantitative	0-100
Conformation	Conformation	Quantitative	0-100
Production	Cow Rating	Quantitative	60-140
Genetics	Genetics	Quantitative	0-100

Table 3.4 - Variables considered in Module Four (obtained from DHI Specialist 2).

MANAGEMENT AREA	VARIABLES	TYPE	ALLOWABLE VALUES
Health	Irrecoverable Diseases	Logical	yes/no
	Feet & Leg Problems	Logical	yes/no
	Total Number of Inseminations	Quantitative	1-9
	Udder Problems	Logical	yes/no
Production	Cow Rating	Quantitative	60-140
Miscellaneous	Lactation Number	Quantitative	1-20
	Culling Rate	Quantitative	0-99
	Number of Cows in Herd	Quantitative	$1-\infty$

Table 3.5 - Variables considered in Module Five (obtained from Researcher 1).

MANAGEMENT AREA	VARIABLE	TYPE	ALLOWABLE VALUES
Production	305-day Milk Production (All Lactations)	Quantitative	2000-25000
	305-day Fat Production (All Lactations)	Quantitative	10-1000
	305-day Protein Production (All Lactations)	Quantitative	10-1000
Production/ Reproduction	Lactation Number	Quantitative	1-20
Reproduction	Avg. Calving Interval	Quantitative	250-750
	Avg. Days to First Service	Quantitative	1-365
	Avg. Total Number of Inseminations	Quantitative	1-10
Miscellaneous	Other Costs	Quantitative	1-1000

Table 3.6 - Variables considered in Module Six (obtained from Researcher 2).

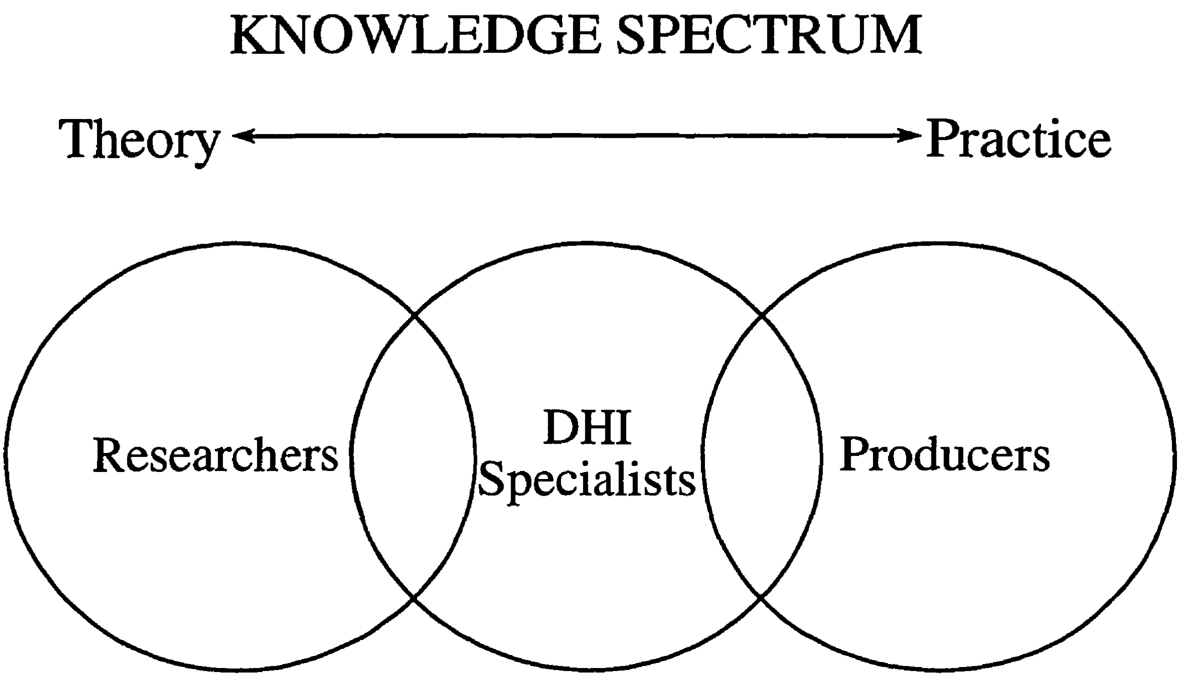

Figure 3.1 - Expert-type knowledge spectrum.

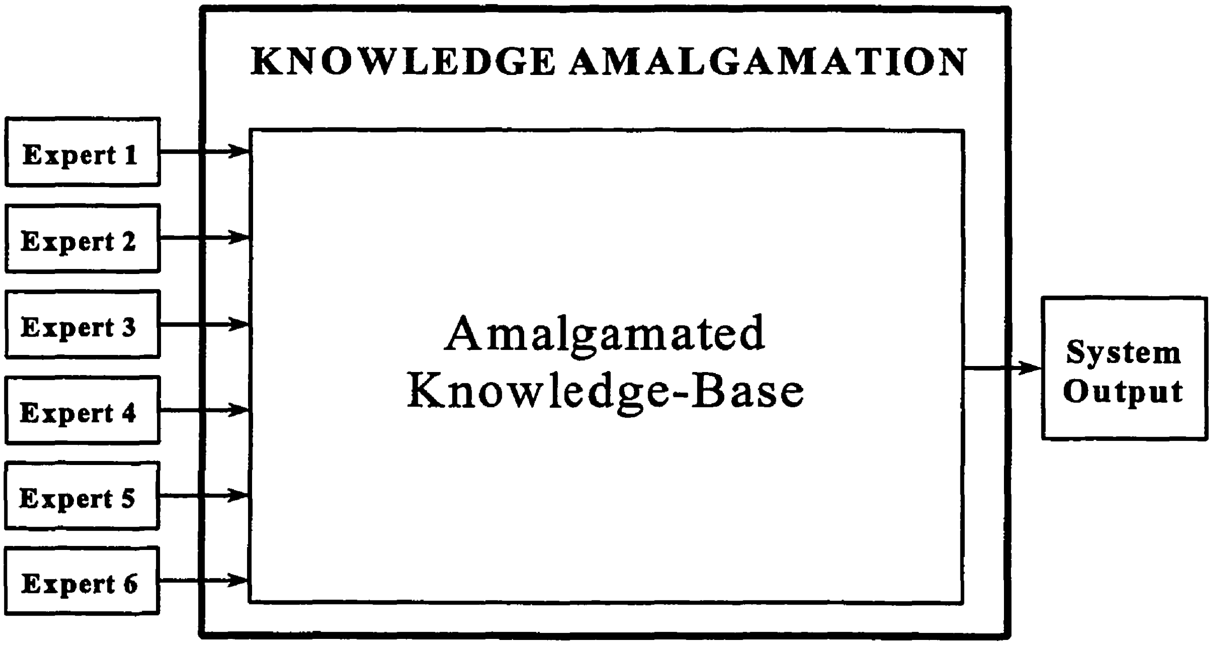

Figure 3.2 - An example of an amalgamated DSS.

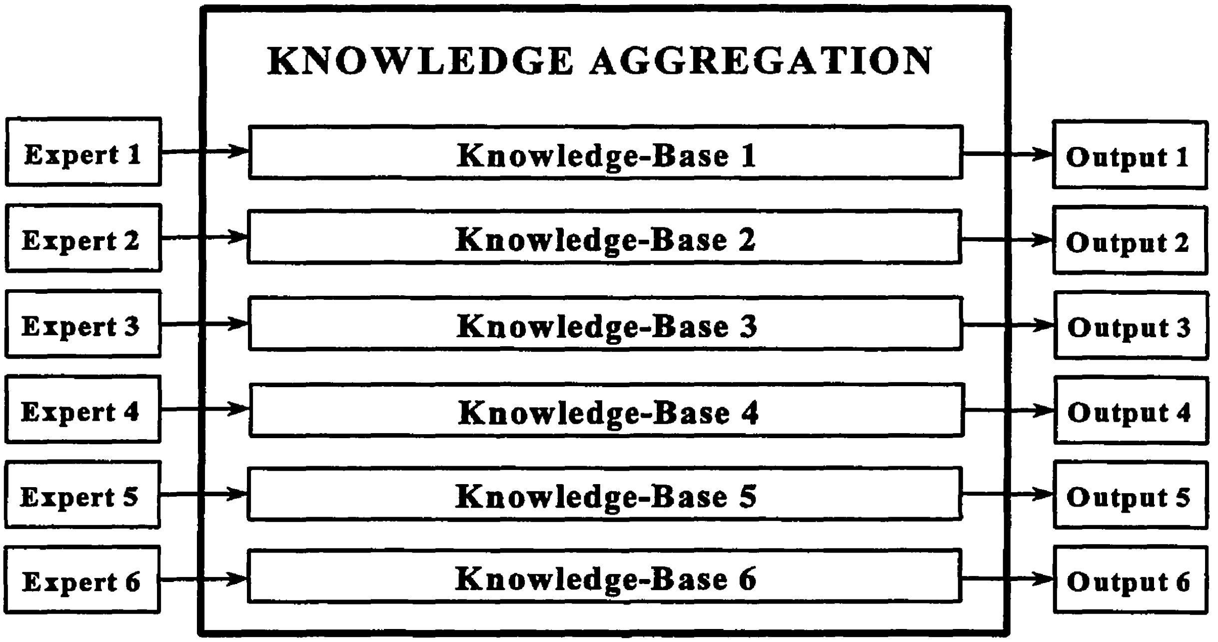

Figure 3.3 - An example of an aggregated DSS.

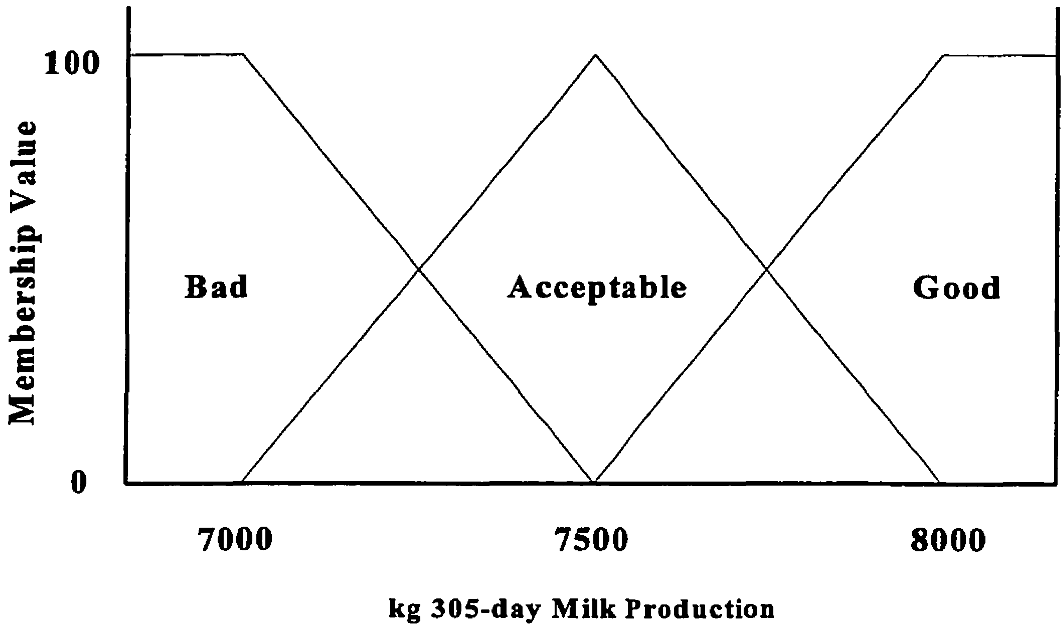

Figure 3.4 - Sample fuzzy sets for 305-day milk production used in Module One for first lactation heifers.

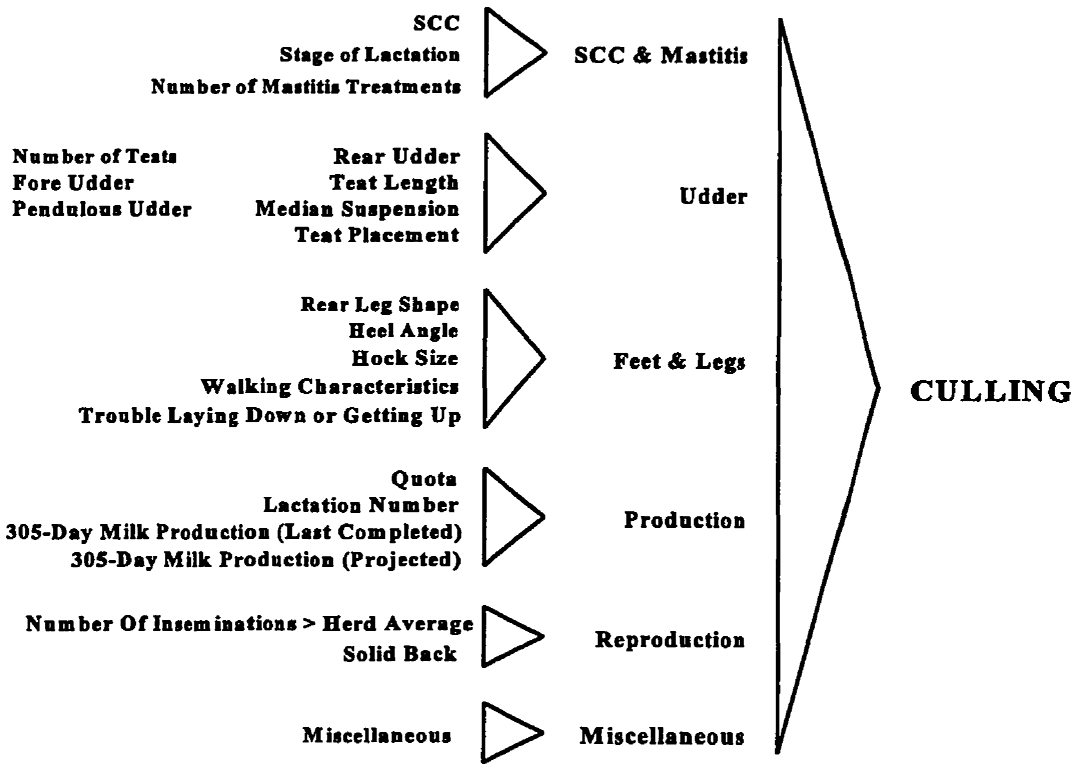

Figure 3.5 - Module One decision tree representing the decision-making process of Producer 1.

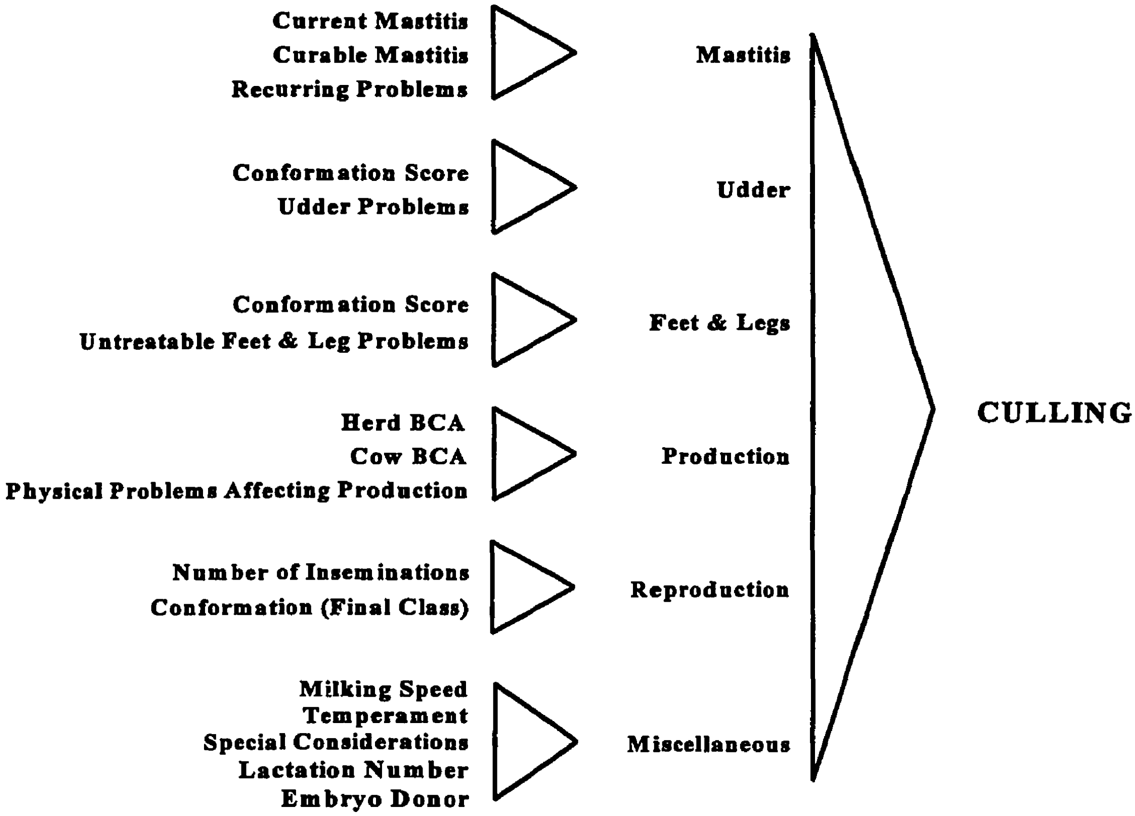

Figure 3.6 - Module Two decision tree representing the decision-making process of Producer 2.

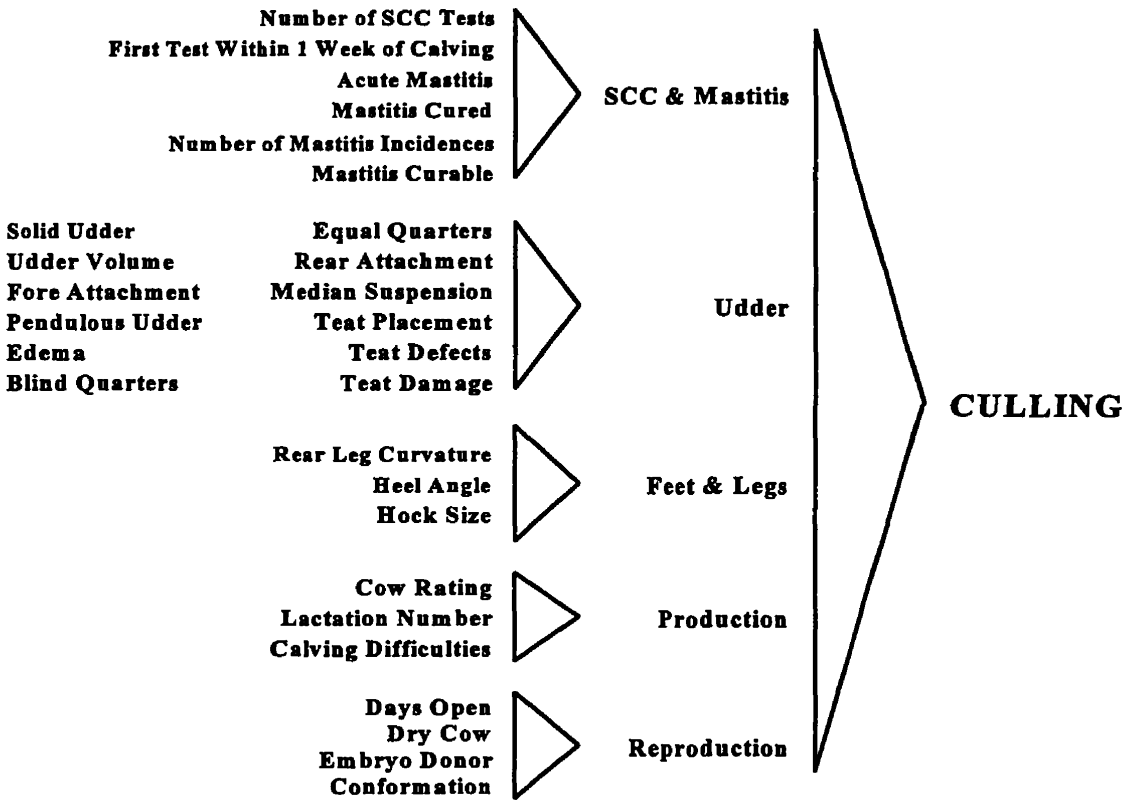

Figure 3.7 - Module Three decision tree representing the decision-making process of DHI Specialist 1.

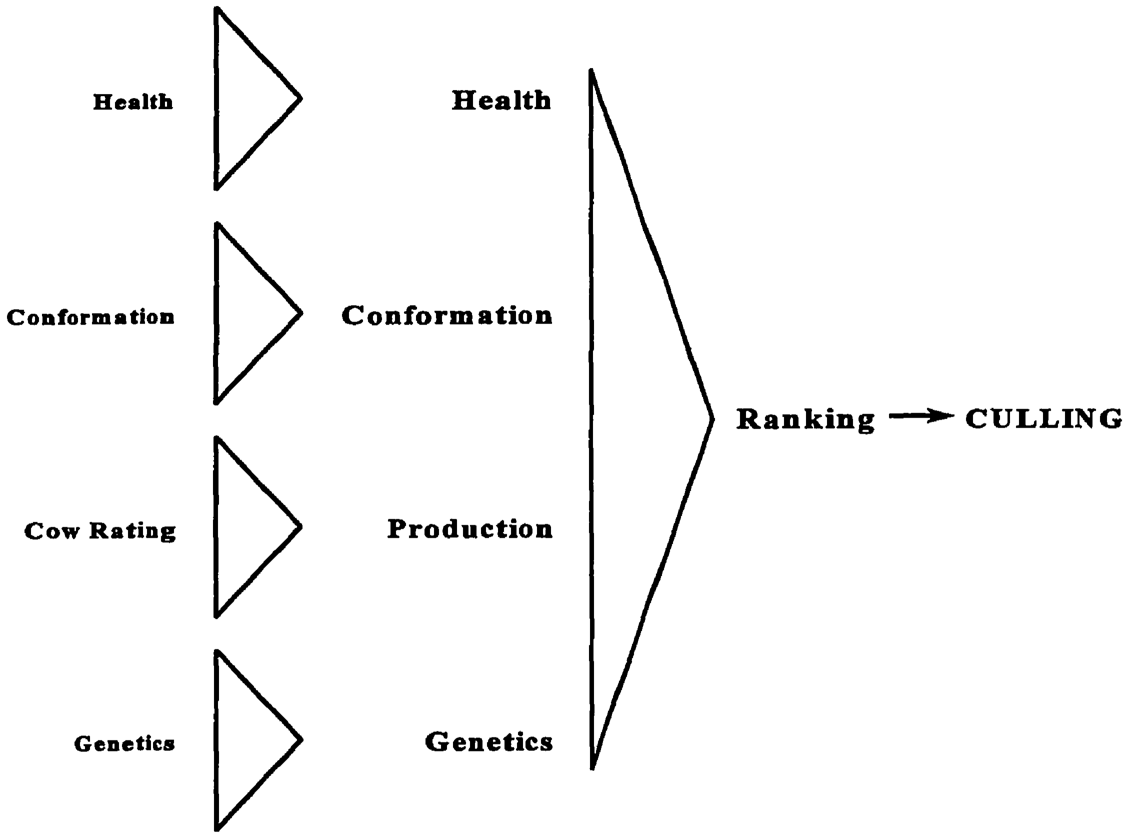

Figure 3.8 - Module Four decision tree representing the decision-making process of DHI Specialist 2.

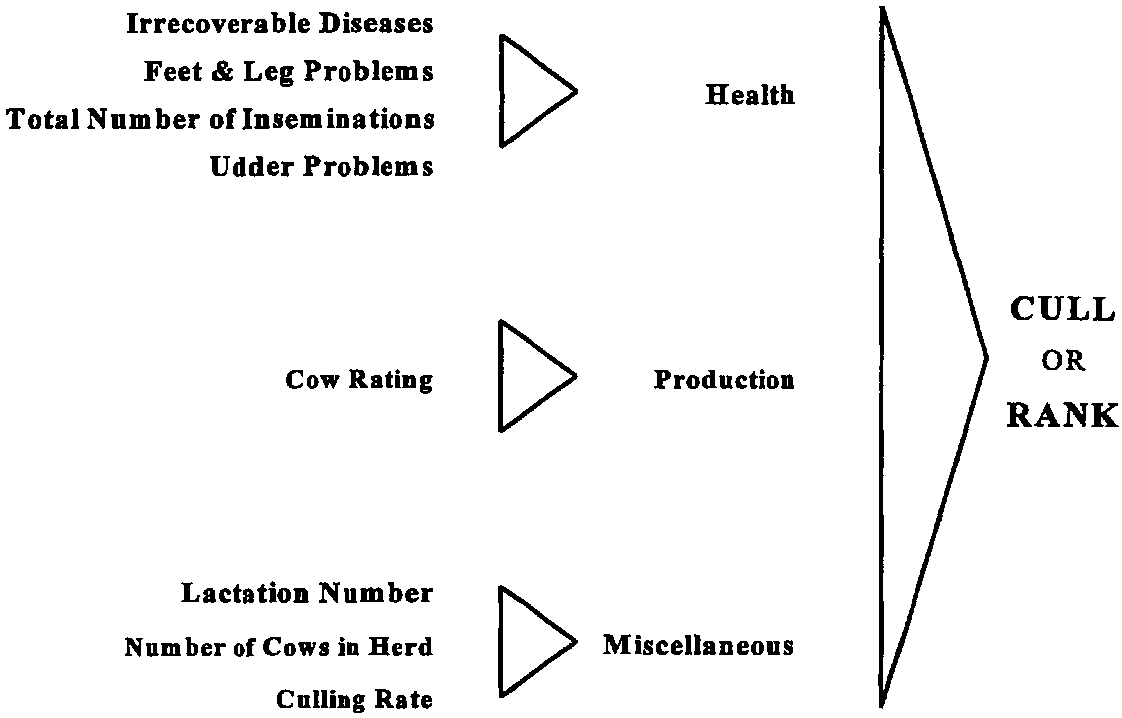

Figure 3.9 - Module Five decision tree representing the decision-making process of Researcher 1.

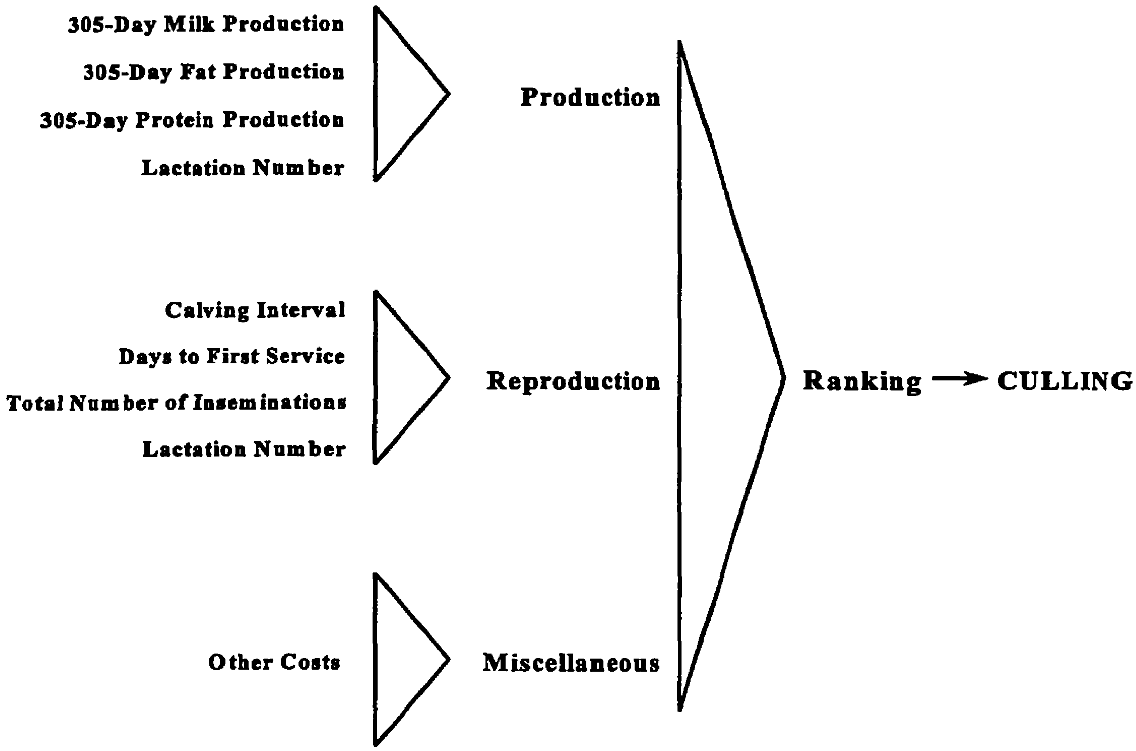

Figure 3.10 - Module Six decision tree representing the decision-making process of Researcher 2.

CONCLUSIONS

The results obtained from the development of a first-generation prototype decision-support system demonstrates that fuzzy logic permits the encoding of knowledge using a terminology which is close to that used by experts, i.e., based on linguistic descriptions. This eases the encoding of knowledge, which may accelerate the development and the implementation of DSSs that accurately reproduce the reasoning of specialists. This would make them more rapidly available to producers. Since fuzzy logic seems promising in the development of knowledge-based systems, it will be used to develop a more complete DSS aimed at helping dairy producers to establish their herd breeding policy, their herd breeding programme and their breeding programme for individual cows.

Through the development of the first-generation prototype, it has been demonstrated that domain experts differ in their use of language and this must be considered when developing decision-support systems. If the knowledge of more than one domain experts is acquired for implementation into a fuzzy DSS, the knowledge can be amalgamated or aggregated. The latter approach ensures that the knowledge of each expert is maintained independently in his/her own personal language and would seem to allow more flexibility with regard to end-user applications.

In constructing a second-generation prototype, six modules were developed with the objective of recommending culling decisions. Although six functional modules have been developed and are capable of recommending culling decisions independently, it was envisaged that these modules would be combined in order to develop a single aggregated decision support system. In fact, this has proven difficult, due, in part, to a number of factors. Firstly, there exists a significant amount of "overlap" between the considerations of the experts, requiring the maintenance of a significant amount of variables. Since the conclusions of each module are not equivalent, it would be difficult to combine the modules. Finally, validation of this DSS would be extremely difficult because the intermediate objectives of the experts are not the same and, therefore, the culling decisions are not expected to be uniform either. A more appropriate means of validation might, therefore, be field testing which would be undertaken by the potential users of the DSS.

For these reasons, future decision support systems should limit the acquisition of expert knowledge to fewer experts, or limit the amount of "overlap" permitted between experts. The aggregated approach has many advantages but requires greater uniformity between modules developed. Finally, the applicability of integrating fuzzy logic with decision support system development is limited. Many of the variables considered by the experts are unavailable through traditional databases, and many other variables are already described qualitatively. Therefore there is no need for the *translation* which is accomplished through the use of fuzzy logic.

The DSS developed during the course of this research were developed with the expert system package; GURU Version 3.0 (by MDBS). Since this package can be considered reasonably old, by computer technology standards, it only permits the creation of a text-based user interface. Graphics-based packages would be of benefit to DSS construction by allowing for the creation of a more user-friendly interface, for example, an expert system shell combined with a graphics development package, such as Visual Basic.

150mm

6″

APPLIED IMAGE . Inc

1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989