

Kristensen, A. R. 1993. *Markov Decision Programming Techniques Applied to the Animal Replacement Problem*. Dissertation for the dr. agro. degree. Dina KVL, Department of Animal Science and Animal Health, The Royal Veterinary and Agricultural University, Copenhagen. ISBN 87 7026 332 9.

Chapter I. Introduction (p 7-12)

Introduction

In the late fifties Bellman (1957) published a book entitled “*Dynamic Programming*”. In the book he presented the theory of a new numerical method for the solution of *sequential decision problems*. The basic elements of the method are the “*Bellman principle of optimality*” and *functional equations*. The idea may be illustrated as follows.

Consider a system being observed over a finite or infinite time horizon split up into periods or *stages*. At each stage, the *state* of the system is observed, and a *decision* (or an *action*) concerning the system has to be made. The decision influences (deterministically or stochastically) the state to be observed at the next stage, and depending on the state and the decision made, an immediate *reward* is gained. The expected total rewards from the present stage until the end of the planning horizon is expressed by a *value function*. The relation between the value function at the present stage and the one at the following stage is expressed by the *functional equation*. Optimal decisions depending on stage and state are determined backwards step by step as those maximizing the right hand side of the functional equation. This way of determining an optimal *policy* is based on the Bellman principle of optimality which says: “An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision” (Bellman, 1957 p. 83).

During the following years Bellman published several books on the subject (Bellman, 1961; Bellman and Dreyfus, 1962; Bellman and Kalaba, 1965). The books were very enthusiastic, and the method was expected to be the solution to a very wide range of decision problems of the real world. The expectations were so great, and they were adduced with such a conviction, that Johnston (1965) ironically compared dynamic programming to a new religion. Others regarded the method as a rather trivial computational device.

Similar stories might be told regarding other new numerical methods, as for instance linear pro-

gramming. However, after some years, the applicational scopes of the methods are encircled. Most often the conclusion is that the method is neither an all-embracing technique nor a triviality. Between these extremities a rather narrow range of problems remains where it is a powerful tool. Other problems are, in some cases, not suitable to be solved by the method. In other cases alternative methods are better.

This also turned out to be the case in dynamic programming. One of the basic elements of dynamic programming is the sequential approach, which means that it fits sequential decision problems best. An obvious example of a sequential decision problem is *the replacement problem*. If an asset is used in a production process it is relevant to consider at regular time intervals whether the present asset should be replaced or it should be kept for an additional period. Thus dynamic programming is a relevant tool, but if the traits of the asset are well defined and their precise behavior over time is known in advance, there are other methods that might be applied to determine the optimal replacement time analytically. On the other hand, if the traits of the asset are affected by *random* variation over time and among assets (as it is the case when the asset is an animal), the replacement decision will depend on the present observations of the traits. In that case dynamic programming is an obvious technique to be used in the determination of an optimal replacement policy.

Having identified dynamic programming as a relevant method to be used with the animal replacement problem, we shall continue on the historical development. In 1960 Howard published a book on “*Dynamic Programming and Markov Processes*”. As will appear from the title, the idea of the book was to combine the dynamic programming technique with the mathematically well established notion of a *Markov chain*. A natural consequence of the combination was to use the term *Markov decision process* to describe the notion. Howard (1960) also contributed to the solution of infinite stage problems, where the *policy itera-*

tion method was created as an alternative to the stepwise backward contraction method, which Howard called *value iteration*. The policy iteration was a result of the application of the Markov chain environment and it was an important contribution to the development of optimization techniques.

The policy iteration technique was developed for two criteria of optimality, namely maximization of total expected *discounted* rewards and maximization of expected *average* rewards per stage. Later on, Jewell (1963) presented a policy iteration technique for the maximization of average rewards over time for *semi-Markov decision processes*, which are Markov decision processes of which the stage length is a random variable. Howard (1971) presented a value iteration method for semi-Markov decision processes.

For the sake of completeness it should also be mentioned that *linear programming* was early identified as an optimization technique to be applied to Markov decision processes as described by, for instance, Hadley (1964), but no animal replacement models known to the author have applied that technique. This is in accordance with a conclusion of White and White (1989) that policy iteration (except in special cases) is more efficient than linear programming.

Since the publication of the first mentioned book by Howard (1960) an intensive research in Markov decision programming has been carried out. Many results have been achieved concerning the relations between the various optimization techniques and criteria of optimality. Reviews of these developments are given by van der Wal and Wessels (1985) as well as White and White (1989).

Already three years after the book by Howard (1960), an application to the dairy cow replacement problem was published by Jenkins and Halter (1963). Their model only included the trait "lactation number" (at 12 levels), and the permanent value of the study was only to illustrate that Markov decision programming is a possible tool to be applied to the problem. A few years later, however, Giaever (1966) published a study which represents a turning-point in the application of the method to the animal (dairy cow) replacement problem. He considered all three optimization techniques (value iteration, policy iteration and linear programming), described how to ensure that all mathematical conditions were satisfied, and presented an eminent model to describe the production and feed intake

of a dairy cow. The work of Giaever (1966) has not got the credit in literature that it deserves (maybe because it is only available on microfilm). In a review by van Arendonk (1984) it is not even mentioned.

During the following 20 years, several dairy cow replacement models using Markov decision programming were published, but from a *methodological* point of view none of them have contributed anything new compared to Giaever (1966). Several studies, however, have contributed in *other* ways. Smith (1971) showed that the rather small model of Giaever (1966) with 106 states did not represent the upper limit. His state space included more than 15 000 states. Kristensen and Østergaard (1982) as well as van Arendonk (1985; 1986) and van Arendonk and Dijkhuizen (1985) studied the influence of prices and other conditions on the optimal replacement policy. Other studies (Killen and Kearney, 1978; Reenberg, 1979) hardly reached the level of Jenkins and Halter (1963). Even though the sow replacement problem is almost identical to that of dairy cows, very few studies on sows have been published. The only exceptions known to the author are Huirne (1990) and Jørgensen (1992).

A study of Ben-Ari et al. (1983) deserves special attention. As regards methodology it is not remarkable, but in that study the main difficulties of the animal replacement problem were identified and clearly formulated. Three features were mentioned:

- 1) *Uniformity*. The traits of an animal are difficult to define and measure. Furthermore the random variation of each trait is relatively large.
- 2) *Reproductive cycle*. The production of an animal is cyclic. It has to be decided *in which* cycle to replace as well as *when* to replace inside a cycle.
- 3) *Availability*. Only a limited supply of replacements (heifers or gilts) is available.

The first feature in fact covers two different aspects, namely *uniformity* because the traits are difficult to define and measure, and *variability* because the traits vary at random among animals and over time. The third feature is an example of a *herd restraint*, i.e. a restriction that applies to the herd as a whole and not to the individual animal. Other examples of herd restraints are a production quota or

a limited housing capacity. We shall therefore consider the more general problem of herd restraints.

We may conclude that when the research presented in this thesis was initiated, the methodological level concerning the application of Markov decision programming to the animal replacement problem was represented by Giaever (1966). The main difficulties that the method should overcome had been identified by Ben-Ari et al. (1983). If we compare the approach of Giaever (1966) to the difficulties that it ought to solve, we may conclude that the problems related to *variability* are directly solved, and as it has been shown by Kristensen and Østergaard (1982) as well as van Arendonk (1985), the problems concerning the *cyclic production* may readily be solved without any methodological considerations. The only problem concerning variability and cyclic production is that to cover the variability the state variables (traits) have to be represented by many levels, and to deal with the cyclic production a state variable representing the stage of the cycle has to be included. Both aspects contributes significantly to an explosive growth of the state space. We therefore face a *dimensionality* problem. Even though all necessary conditions of a Markov decision process are met, the solution in practice is prohibitive even on modern computers. The problems concerning uniformity and herd restraints are *not* solved by the approach of Giaever (1966).

The purpose of this thesis is to adapt the Markov decision programming techniques to be able to cope with the animal replacement problem in a satisfactory way. The problems to be solved (totally or partially) have been identified as the dimensionality problem, the uniformity problem and the problems caused by herd restraints. A secondary purpose is to illustrate and discuss the applicational perspectives of the techniques. All numerical results of the thesis refer to dairy cows, but recently Markov decision programming has also been applied to sows (Huirne, 1990; Jørgensen, 1992). Since the sow replacement problem does not differ very much from that of dairy cows, the same methodological problems arise, and the results of this thesis are therefore relevant in sow replacement models too.

In Chapter II a systematic survey of the developed techniques is given. In Chapters III-VIII the individual techniques are described in details. The applicational perspectives are discussed in Chapter

IX, and Chapters X and XI are examples of such applications.

References

- Ben-Ari, Y., Amir, I. and Sharar, S. 1983. Operational replacement decision model for dairy herds. *Journal of Dairy Science* 66, 1747-1759.
- Bellman, R. E. 1957. *Dynamic Programming*. Princeton: Princeton University Press.
- Bellman, R. E. 1961. *Adaptive control processes: a guided tour*. Princeton: Princeton University Press.
- Bellman, R. E. and Dreyfus, S. E. 1962. *Applied dynamic programming*. Princeton: Princeton University Press.
- Bellman, R. E. and Kalaba, R. 1965. *Dynamic programming and modern control theory*. New York: Academic Press.
- Giaever, H. B. 1966. *Optimal dairy cow replacement policies*. Ph.D. thesis. Ann Arbor, Michigan: University Microfilms, Inc.
- Hadley, G. 1964. *Nonlinear and dynamic programming*. Reading, Massachusetts: Addison-Wesley.
- Howard, R. A. 1960. *Dynamic programming and Markov processes*. Cambridge, Massachusetts: The M. I. T. Press.
- Howard, R. A. 1971. *Dynamic probabilistic systems. Volume II: Semi-Markov and decision processes*. New York: John Wiley & Sons, Inc.
- Huirne, R. B. M. 1990. *Computerized management support for swine breeding farms*. Ph.D. thesis. Wageningen Agricultural University.
- Jenkins, K. B. and Halter, A. N. 1963. A multi-stage replacement decision model. *Agricultural Experiment Station, Technical Bulletin 67*, Oregon State University.
- Jewell, W. 1963. Markov renewal programming I and II. *Operations Research* 11, 938-971.
- Johnston, J. 1965. Dynamic programming and the theory of the firm. *Journal of Agricultural Economics* 16, 532-543.
- Jørgensen, E. 1992. Sow replacement: Reduction of state space in dynamic programming model and evaluation of benefit from using the model. *Livestock Production Science*. Submitted for publication.
- Killen, L. and Kearney, B. 1978. Optimal dairy cow replacement policy. *Irish Journal of Agricultural Economics and Rural Sociology* 7, 33-40.
- Kristensen, A. R. and Østergaard, V. 1982. Optimalt udskiftningsstidspunkt for malkekøer fastlagt ved en stokastisk model. *Beretning fra Statens Husdyrbrugsforsøg* 533. In Danish with English Summary.
- Reenberg, H. 1979. Udskiftning af malkekøer. En stokastisk udskiftningsmodel. *Jordbrugsøkonomisk Institut Memorandum 6*. In Danish with English Summary.
- Smith, B. J. 1971. The dairy cow replacement problem: An application of dynamic programming. *Florida Agricultural Experiment Station Bulletin 745*. Gainesville, Florida.
- van Arendonk, J. A. M. 1984. Studies on the dairy cow replacement policies. I. Evaluation of techniques to determine the optimum time for replacement and to rank cows on future profitabilities. *Zeitschrift für Tierzüchtung und Züchtungsbiologie* 101, 330-340.

- van Arendonk, J. A. M. 1985. Studies on the dairy cow replacement policies. II. Optimum policy and influence of changes in production and prices. *Livestock Production Science* 13, 101-121.
- van Arendonk, J. A. M. 1986. Studies on the dairy cow replacement policies. IV. Influence of seasonal variation in performance and prices. *Livestock Production Science* 14, 15-28.
- van Arendonk, J. A. M. and Dijkhuizen, A. A. 1985. Studies on the dairy cow replacement policies. III. Influence of variation in reproduction and production. *Livestock Production Science* 13, 333-349.
- van der Wal, J. and Wessels, J. 1985. Markov decision processes. *Statistica Neerlandica* 39, 219-233.
- White, C. C. III and White, D. J. 1989. Markov decision processes. *European Journal of Operational Research* 39, 1-16.