J. Opl Res. Soc. Vol. 44, No. 11, pp. 1073-1096
Printed in Great Britain. All rights reserved
0160-5682/93 $9.00+0.00

A Survey of Applications of Markov Decision Processes

D. J. WHITE
Department of Decision Theory, University of Manchester

A collection of papers on the application of Markov decision processes is surveyed and classified according to the use of real life data, structural results and special computational schemes. Observations are made about various features of the applications.

Key words: Markov decision processes, applications

INTRODUCTION

In White[1] a survey of 'real' applications of Markov decision processes was presented where 'real' means studies where the results were actually implemented or at least had an effect on the actual decisions taken. In a later paper[2], a slightly more liberal interpretation was allowed in order to include studies of actual problems, which were based on real data, and where the object was to help solve these actual problems.

In the process of acquiring that material, other material of a more hypothetical nature was acquired in which, in most cases, the authors modelled problem situations which did not represent actual problem situations. However, these models were almost certainly motivated by the author's knowledge of related real situations. This paper presents a brief description of this work, as well as highlighting certain aspects of these papers. Not all of these papers are completely hypothetical and, indeed, some of them do use real data. Where this is the case, an appropriate comment is made.

The survey is restricted to published papers and to working papers which exhibit a substantial content for the problem being studied, and does not include material which occurs, in abundance, as illustrations in books, with two exceptions.

The survey is not intended to be anything more than a random survey and, in most instances, can in no way reflect the substantial body of publications which certainly exist in many of the areas covered. For example, there is much more work on modelling queueing situations as Markov decision processes than is discussed in this paper, and similarly with replacement, maintenance and repair, and so on. But this is not true for all. For example, the applications of Markov decision processes to motor insurance claims is, as yet, not a large area.

There is, then, the question of what useful purposes such a limited survey may serve. The following purposes are relevant, namely:

(i) to provide a source of much more substantial applications material even though somewhat hypothetical, which may strengthen teaching programmes in the general area of Markov decision processes;
(ii) to motivate academics to produce comprehensive surveys within the specific areas of application;
(iii) to motivate practitioners towards thinking, in their respective areas, about the possibility of using Markov decision processes;
(iv) to learn a little from the special features of the specific papers and to suggest possible research questions.

In the survey comments, some attention will be given to the last point.

Correspondence: D. J. White, Faculty of Economic and Social Studies, Department of Decision Theory, University of Manchester, Manchester M13 9PL, UK

THE SURVEY

The survey is based on papers which the author had in his possession at the time of writing. It does not include the 40 papers cited in White[1] nor a further 17 papers cited in White[2] where, in the latter, the papers are rather closer to the real applications spirit of Reference 1.

The papers are categorized into 18 areas, although there is inevitably some choice as to which category should be assigned to some papers, e.g. the paper by Dreyfus[3], which has both a replacement and a production content. Fortunately, this ambiguity is minimal in this survey.

The 18 areas are given in Table 1. The numbers in the brackets are the numbers of papers cited. The 18th area is nominally an area containing general applications, some of which will be in the other areas.

TABLE 1. *Application areas*

1	Population harvesting	(5)
2	Agriculture	(5)
3	Water resources	(15)
4	Inspection, maintenance and repair	(18)
5	Purchasing, inventory and production	(14)
6	Finance and investment	(9)
7	Queues	(6)
8	Sales promotion	(4)
9	Search	(3)
10	Motor insurance claims	(2)
11	Overbooking	(5)
12	Epidemics	(2)
13	Credit	(2)
14	Sports	(2)
15	Patient admissions	(1)
16	Location	(1)
17	Design of experiments	(1)
18	General	(5)

Within each area, the papers are listed chronologically, except where an author makes several contributions over several years.

The brief details given in the survey are split into four columns, as indicated in Table 2.

TABLE 2. *Brief details*

1	References
2	Short summary of the problem
3	Objective function
4	Comments

The short summary of the problem gives some idea of the decisions to be made, of the context of the problem, and of the state description used.

The objective functions are all expected, discounted or non-discounted, costs or rewards over a finite horizon, or costs or rewards per unit time, over an infinite horizon, sometimes subject to side constraints.

Comments are then made on the use of real data, on the nature of the model, and on particular features such as special policy results.

In making comments about the model, the term 'stochastic dynamic program' is used, as distinct from 'Markov decision process'. The models are all Markov decision process models, but not all of them use functional stochastic dynamic programming equations. Some use equivalent linear programming formulations, although these are in the minority.

The main survey is given in Table 3.

TABLE 3. *Applications of Markov decision processes*

Reference	Short summary of the problem	Objective function	Comments
	1. Population harvesting		
Mendelssohn[4-6]	Decisions have to be made each year as to how many members of a population (e.g. fish) have to be left to breed for the next year. The states are the numbers of the species in each age category. The new states, in the next year, depend on the numbers left to spawn and on a random factor, as well as on the decisions taken.	Expected discounted return over a finite number of years.	Various return functions are used. Real data on fish population are used, but return functions are hypothetical. A standard finite-period stochastic dynamic programming value iteration procedure is used in the 1978 and 1980 papers[4,5]. The particular state transition structure allows the problem to be decomposed into several independent simpler optimization problems[4,5]. In the 1978 and 1980 papers[4,5], given certain conditions, it is shown that certain structural policies are optimal.
Mann[7]	The problem class is similar to that of Mendelssohn[4-6]. The states are the numbers of males and females in the population at the decision epoch, plus an extrinsic factor which influences the growth rates of the population. The extrinsic factor follows a Markov pattern.	As for Mendelssohn[4-6]	A stochastic dynamic programming value iteration procedure was used. Under certain conditions it shows that certain structural policies are optimal.
Ben-Ari and Gal[8]	The problem class is similar to that of Mendelssohn[4-6] The states of the system are the numbers of living animals in a certain category, where a category is a combination of age, expected milk yield and weight.	As for Mendelssohn[4-6]	Real data are used for a Kibbutz Yagur herd. A finite-period stochastic dynamic programming value iteration procedure is used, using functional approximation to overcome the state-space cardinality problem.
	2. Agriculture		
Brown *et al.*[9]	Each year decisions have to be made as to whether or not a given area should be planted or fallowed. The states of the system are the soil moisture content and the probability forecast that the current year's growing season precipitation will be in the lowest category, at the decision epoch. The new states in the next year depend on the actual precipitation in the current year and on the decisions taken.	Expected discounted return over a finite number of years.	Real data are used for the Havre, Montana, Willington, North Dakota regions. A standard finite stochastic dynamic programming value iteration is used. The study is used for evaluating the economic value of the forecasting system.
Onstad and Rabbinge[10]	Throughout a season decisions have to be made as to whether or not treatment should be applied to protect a crop against pests. The states of the system are surrogate measures of pest level. The new states at the next decision epoch depend upon rainfall and on the decisions taken.	Expected cost throughout the season, inclusive of the effect of pests on the harvest level.	Real data are used. The problem is formulated as a finite-horizon stochastic dynamic program and solved using successive approximations.
Jacquette[11], Conway[12], Feldman and Curry[13]			These papers contain many references to stochastic dynamic programming models for pest control.

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
		3. Water resources	
Little[14]	At each decision epoch decisions have to be made as to how much water is to be used to generate electrical power, when alternative methods of power generation exist, and demands are specified over a given time period. The states of the system are the current water levels and inflows in the previous period. The states at the next decision epoch depend upon previous and current inflows as well as on the decisions taken.	Expected cost over a finite horizon, where the cost in any period includes the cost of alternative energy generation to meet demand.	Uses real data for the Grand Conlee plant on the Columbia River. Standard finite-period stochastic dynamic programming is used.
Buras[15]	This deals with a serial two-reservoir system, in which decisions have to be made as to how much water to divert from the higher level reservoir to the lower level reservoir, and how much to release from each for irrigation purposes. The states at each decision epoch are the current reservoir levels and the current flow between them. The states at the next decision epoch depend upon the flow into the higher level reservoir as well as on the decisions taken.	Expected discounted profit over a finite horizon arising from subsequent use of irrigation water.	Uses standard finite-period stochastic dynamic programming.
Su and Deininger[16]	Decisions have to be made each month as to how much water to release from a lake. The states are the current lake levels, and either the previous month's inflow or the current predicted inflow, if accurate. The new states at the next decision epoch depend on the current inflow as well as on the decisions taken.	Expected loss per unit time over an infinite horizon and the expected infinite horizon discounted loss, where the loss reflects losses to shore property and to navigation.	The problem exhibits annual cyclical behaviour and a modification of the stochastic dynamic programming relative-value iteration procedure is used to solve it.
Russell[17]	The problem is similar to that of Little[14], except that the states are simply the current reservoir levels, and the impact of water levels on matters other than just power generation is considered.	Expected discounted cost over a finite horizon, where the costs include costs arising from water levels being too high (floods) or too low (amenity).	Standard finite-period stochastic dynamic programming is used and some analytic properties of an optimal solution derived.
Arunkumar and Chon[18]	This problem is similar to that of Buras[15], except that random inputs are allowed into both reservoirs and the state description does not include the current release from the higher level reservoir to the lower level one.	Expected cost over a finite time horizon, where the costs reflect irrigation and recreational costs.	Real data for the Shasta–Folsam parallel system over 61 years are used. Finite period stochastic dynamic programming is used, incorporating the annual cyclicity of the problem, and using structural policies, but without validation.
Shamir[19]			This paper contains references to stochastic dynamic programming models for water resource management.

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
Turgeon[20-22]	The 1980 paper[20] is an extension of the problem of Little[14] where power generation is the only objective, but where several reservoirs exist and decisions have to be made as to how much power should be generated from each, and the states include only current reservoir levels. Decisions are taken monthly.	Expected cost over a finite time horizon, inclusive of the negative cost of selling excess power and the cost of making up deficits.	Real data are used for the Quebec River. Standard finite period stochastic dynamic programming is used in conjunction with a special approximating decomposition method necessitated by the high cardinality of the problem.
	The 1981 paper[21] is similar to the 1980 paper[20], except that the reservoirs are in series and the decisions are the releases down stream (as in the case of Arunkumar and Chon[18]) as well.	The expected reward over a finite time horizon where the rewards are specified as some functions of the reservoir levels.	Real data are used for a four-reservoir system. Finite-period stochastic dynamic programming is used, incorporating a special aggregation procedure necessitated by the high cardinality of the problem.
	The 1985 paper[22] is similar to the 1980 paper[20], except that it has weekly decision epochs, and is subject to constraints on the frequency with which critical reservoir levels are exceeded.	As for the 1980 paper[20], but subject to the constraints specified in the preceding column.	Uses finite-period stochastic dynamic programming, using Lagrange multipliers to incorporate the constraints into the objective function. The structure of the problem allows a reservoir aggregation procedure to be used to give an optimal policy.
Krzysztofowicz and Jagannathan[23]	Over a period of time decisions have to be made about the release of water from a single reservoir, to minimize the effect of flooding and, at the same time, provide enough water for irrigation purposes later in the year. The states at each decision epoch are the current reservoir levels and the previous period's inflows. The new states at the next decision epoch depend upon the current inflow as well as on the decisions taken.	Expected utility per year of the flooding effects up to the irrigation period and of the irrigation value of water released up to that time.	Real data used for an Indian monsoon period. The problem is formulated as an infinite-horizon stochastic dynamic program. Uses a successive approximations method to solve the infinite horizon problem.
Yakowitz[24]			This paper contains references to stochastic dynamic programming models for water resource management.
Yeh[25]			As for Yakowitz[24].
Sarin and El Benni[26]	A pumping system pumps water from wells to consumers and to a reservoir system. The decisions are the various combinations of pump operations to supply demands. Decisions are taken each four hours.	Expected pumping costs over a week.	No model is given, but real data for Lancaster, Ohio, used. A finite-horizon stochastic dynamic programming approach is used.
Stedinger *et al.*[27]	The problem is similar to that of Su and Deininger[16], with the exception that power production, flooding, and irrigation targets are set. The states are the reservoir levels, and one of several future	Expected weighted squared excess, or deficit, of flood levels, irrigation provisions and power provisions, from specified targets, over a finite horizon.	Real data for the Aswan dam in the Nile River Basin are used. Standard finite period stochastic dynamic programming is used. Simulations indicated that the use of best forecasts for future

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
	hydrological predictions are used, inclusive of the previous period's inflow and a best forecast of inflow.		hydrological inflows produced the best policy.
Krzysztofowicz[28]	Decisions have to be made concerning the commitment to supplying water for various purposes at a later date. The states at each decision epoch are the (sure) supplies of water available and the forecasts of water to be available at the later required dates. The states at the next decision epoch depend upon current forecasts and upon the decisions taken.	Expected utility of the actual water available for the required purposes, where the utility is a function of the time when the decision is made to commit the suppliers to providing specific water supplies.	Real data for the Twin Springs and Weiser forecasting systems and for the Owyhee Irrigation Project are used. Uses finite period stochastic dynamic programming.
	4. *Inspection, maintenance and repair*		
Dreyfus[3]	Decisions have to be made as to which bladder producing tyres should be used for production of a tyre, and which should be replaced. The states of the system are the numbers of tyres produced to date on each bladder and the residual number of tyres required. The new states at the next decision epoch depend upon whether or not bladders fail in the current production operation as well as on the decisions taken.	The expected cost of eventually producing the required tyres.	The problem is modelled as an absorbing-state stochastic dynamic program, and computations work backwards from zero requirements to N requirements.
Gluss[29]	Decisions have to be made as to which module, in a multi-module system, should be tested, and then which component should be tested, when the system has developed a fault. The states of the system are the current probabilities of the fault residing in any module, or in any component. The new states at the next decision epoch depend on information obtained at the current test and on the decisions taken.	Expected time or cost to locate the fault.	The problem is modelled as an absorbing-state stochastic dynamic program. Some structural results are obtained concerning the nature of the optimal policy. Two models are considered, the first allowing a module to be tested as a whole, and the second allowing only component testing.
White[30]	Decisions have to be made as to when to carry out maintenance to restore the operational performance of a deteriorating production process. The states of the system at any decision epoch are the known parameters which govern the costs for any production run length. These states change in a probabilistic manner when maintenance is carried out.	The expected production cost per unit time over an infinite horizon.	The problem is modelled as an infinite-horizon semi-Markov decision process. Under certain conditions structural optimal policies for the determination of the next maintenance epoch are given.
White[31]	When a system fails, decisions have to be made as to the category of component with which to replace a failed component, bearing in mind	Expected replacement costs until the system fails completely.	The problem is modelled as an absorbing state stochastic dynamic program. The solution procedure is a backtracking one, beginning

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
	the probabilistic residual life of the system. The states of the system are its ages. The new states of the system are its ages when a component fails again.		with the system's maximal age. Continuous and discrete time problems are considered.
White[32]	This problem has similarities to that of White[30]. However, decisions may be made to inspect and to carry out maintenance on the basis of the inspection result.	Expected cost per unit time over an infinite horizon, where the costs include operations, inspection and maintenance costs.	The problem is modelled as an infinite-horizon semi-Markov decision process. Policy space is suggested as a method for solving the problem.
Eckles[33]	Decisions have to be made as to whether or not replacements of parts should be made. The true condition of the system is unknown. The states of the system are the age of the system since the last replacement and the complete histories of actions and of observational information, and the latter is reduced to equivalent probability distributions on the system's true conditions. The new states at the next decision epoch depend on new observational information and on the decisions taken.	Expected discounted cost over an infinite time horizon.	The problem is modelled as an infinite-horizon stochastic dynamic program. Successive approximation is suggested as a solution procedure.
Hinomoto[34]	Decisions have to be made concerning the maintenance of a number of interacting activities. The states of the system are the deterioration levels of each of the activities. The new states at the next decision epoch depend upon certain probabilistic factors and upon the decisions taken.	Expected reward per unit time over an infinite time horizon.	The problem is modelled as a linear programming problem. Two models are considered; namely, one in which all activities are treated in a specified sequence, and one in which freedom exists to inspect and maintain any activity in the light of the system's states.
Kao[35]	This problem is similar to that of White[31] except that only replace/do not replace decisions are allowed, and the next decision epoch is determined when there is a change in state.	Expected discounted cost over an infinite horizon and the expected cost per unit time over an infinite horizon.	The problem is modelled as an infinite-horizon stochastic dynamic program. Two models are considered, one of which takes account of the age of the system. Policy space is suggested as a method of solution.
Duncan and Scholnick[36]	This problem has similarities to that of Dreyfus[3]. The states of the system include the component's ages as well as the number of units it has produced. In addition, the problem takes into account the fact that the opportunities for replacement are probabilistic.	Expected replacement cost per unit time over an infinite horizon.	The problem is modelled as an infinite-horizon stochastic dynamic program. Under certain conditions, structural optimal policies for the replacement decision are obtained. Some special cases are considered.
Crabhill[37]	Decisions have to be made about service rates in a multi-machine system subject to breakdown. The states of the system at any decision epoch	The expected cost per unit time over an infinite horizon, where the cost is the sum of the repair cost and the lost production cost.	The problem is modelled as a continuous-time stochastic dynamic program. Some repair rates can be eliminated as being non-optimal for any

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
	are the numbers of machines in working order. The new states depend upon further breakdowns and repairs taking place and on the decisions taken.		state. Conditions are given for the existence of a monotone optimal policy.
Butler[38]	Decisions have to be made as to whether or not a system should be inspected, where an inspection can impair the life of the system (e.g. X-rays). The states of the system are failed, partially failed, and the number of periods which elapsed since the last inspection, if this was satisfactory. The new states at the next decision epoch depend upon the result of the inspection or upon failure and on the decisions taken.	The expected life of the system.	The problem is modelled as an absorbing-state stochastic dynamic program. A finite horizon problem is also considered. Gives some conditions under which it is always optimal to inspect or never optimal to inspect.
Stengos and Thomas[39]	Decisions have to be made as to whether or not any of two pieces of equipment should be taken out of service and overhauled, before they actually fail. The new states depend upon any repairs or failures being completed, and on the decisions taken.	Expected cost per unit time over an infinite horizon where the cost is the cost of equipment being out of service.	The problem is modelled as an infinite-horizon stochastic dynamic program. The relative value algorithm is used and it is shown that structural optimal policies exist.
Sengupta[40]	This paper has some similarities to that of Eckles[33] Decisions have to be made as to whether or not the system should be inspected. The states of the system are the probabilities of the system being in a failed state. The new states depend upon information obtained on any inspection and on the decisions taken.	Expected cost over a finite time horizon, where the costs are the costs of operating in a failed state plus inspection costs.	The problem is modelled as a finite-horizon stochastic dynamic program. Some structural policy results are obtained. An application to medical diagnosis is given.
Hayre[41]	A system deteriorates and decisions have to be made as to whether it should be repaired or replaced at a specified deterioration level. The states of the system are the numbers of repairs carried out on the current system. The new states will be the states when the system again reaches the critical deterioration level.	Expected cost per unit time for an infinite horizon process, where the costs are repair and replacement costs.	The problem is modelled as an infinite-horizon semi-Markov stochastic dynamic program. Under certain conditions a structural optimal policy is obtained.
Tijms and van der Duyn Schouten[42]	A system deteriorates and decisions have to be made as to whether to leave it alone, inspect, or repair, if an inspection has been made and when an inspection has not been made. Decisions are only allowed at a malfunction. The states of the system are the deterioration conditions and the numbers of periods elapsed since this condition was known. The new states	Expected cost per unit time for an infinite horizon process.	The problem is modelled as an infinite-horizon semi-Markov stochastic dynamic program. Policy space is used to get a solution which is structural. The method works well, but the authors say they cannot justify it theoretically.

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
	depend upon the time to the next inspection and its new condition then, as well as on the decisions taken.		
Ohnishi *et al.*[43]	This is similar to the problem of Eckles[33]. The decisions are either to continue with no inspection, to inspect and operate the system, or to replace with a new system. The states of the system are the probabilities of the system being in various deterioration conditions. The new states at the next decision epoch depend upon information obtained, with or without inspection, about the condition of the system, and upon the decisions taken.	Expected discounted costs over an infinite horizon where the costs are the degradation costs, inspection costs, and replacement costs.	The problem is modelled as an infinite-horizon stochastic dynamic program. Stochastic dominance ideas are used to obtain structural optimal policies.
Thomas[44]	In each of a succession of time periods actions have to be taken to inspect, repair or do nothing to a standby unit. The state of the standby unit is its probability of being useable and the number of periods since installation. The new states depend upon the action taken and the results obtained.	Expected time until a catastrophe occurs or the probability that a catastrophe occurs within a given total time period, where a catastrophe is said to arise if the standby unit is required and is not ready.	Absorbing-state stochastic dynamic programming is used, and structural results obtained.
Gaver *et al.*[45]			This paper contains details of various applications of stochastic dynamic programming.
	5. *Purchasing, inventory and production*		
White[31]	Decisions have to be made concerning the size of production runs in order to meet a target requirement, where production is subject to a random defective content. The states of the system at any decision epoch are the residual requirements. The new states at the next decision epoch depend on the defective production in the current production run and on the decisions taken.	Expected cost arising from at least meeting the initial requirement, where the relevant costs are set-up costs and production costs.	The problem is modelled as an absorbing-state stochastic dynamic program. The method of solution is by backwards calculations, beginning with zero requirements.
Kingsman[46]	Decisions have to be made as to how much of a commodity to purchase at each decision epoch in the light of fluctuating prices and a known demand pattern. The states of the system at each decision epoch are the inventory levels and commodity prices. The new states at the next decision epoch depend upon new prices and the decisions taken.	Expected cost in meeting the demand requirements, where the relevant costs are purchase costs and inventory holding costs.	The problem is modelled as a finite-horizon stochastic dynamic program. An optimal structural policy is derived.
Sobel[47]	Decisions have to be made concerning a work-force level, which remains fixed once chosen, and the production	Expected discounted cost arising from meeting the demand requirements, where the relevant costs are	Stochastic dynamic programming is used for finite-horizon non-stationary and infinite horizon stationary

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
	levels at each decision epoch to meet random demands. The states of the system are the inventory levels at each decision epoch. The new states at the next decision epoch depend upon the current demand levels and upon the decisions taken.	inventory holding costs, production costs, and employment costs. Both finite and infinite horizon problems are covered.	cases, and structural results obtained. For the infinite horizon case it is shown how linear programming may be used to obtain a solution.
Kalymon[48]	This problem is similar to that of Kingsman[46], the essential differences being that demand is uncertain, and the commodity prices depend on the previous price and on the time.	Expected discounted cost over a finite horizon where the relevant costs are the inventory holding costs, shortage costs and purchasing costs.	The problem is formulated as a finite horizon and as an infinite-horizon stochastic dynamic program. Structural optimal policies are derived. In addition, bounds on the optimal policy regions are also derived.
Symonds[49]	Decisions have to be made concerning the quantities of a commodity to order at each decision epoch in the face of random demands. The states at each decision epoch are the inventory levels. The new states at the next decision epoch depend upon the current demand and upon the decisions taken.	Expected cost over a finite horizon, where the relevant costs are inventory holding costs, shortage costs, and purchase costs.	The problem is formulated as a finite-horizon stochastic dynamic program. A special method is developed for solving the problem.
White[50]	Decisions have to be made as to which set of jobs, from those available in a given week, should be done, bearing in mind the arrival of new jobs and the possibility of more efficient scheduling combinations later on. The states of the system are the sets of jobs waiting to be done. The new states at the next decision epoch depend on the new jobs arriving in the current period, and on the decisions taken.	Expected production costs over a finite horizon.	The problem is formulated as a finite-horizon stochastic dynamic program.The tractability of the solution depends on a weak coupling assumption which, in effect, requires that overspill from week to week is small.
Thomas[51]	This is similar to the problem of Symonds[49], the essential difference being that the selling price for the product is also a decision to be taken, and the demand depends upon the price set.	Expected discounted cost over a finite period, where the relevant costs are inventory holding costs, backlog costs, and negative revenue costs.	The problem is formulated as a finite-horizon stochastic dynamic program. A structural optimal policy is conjectured, which turns out to be optimal in all the cases studied, but for which no proof is available.
Snyder[52]	This is similar to the problem of Symonds[49], the essential difference being that the problem is in continuous time, decisions are made when a demand is received, and there is a lead time.	Expected cost over a finite number of inventory reviews, where the costs are as for Symonds[49], except that backlog costs are used instead of shortage costs.	The problem is formulated as a finite-number of decision epochs stochastic dynamic programs. No analysis is given.
Karmarkar[53]	Decisions have to be made as to the levels of various operations required to increase the inventory levels of a range of products required to meet an uncertain demand. The states at each decision epoch are the inventory levels of each of the products. The new	Expected discounted cost over a finite horizon, where the relevant costs are production costs, inventory holding costs, and backlogging costs.	The problem is formulated as a finite-horizon stochastic dynamic program. Under certain conditions structural optimal policies are shown to exist. The problem differs from the usual production–inventory control problem in that the decisions

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
	states at the next decision epoch depend upon the current demands and on the decisions made.		are the physical operations required to produce the requisite inventories.
Parlar[54]	Decisions have to be made as to how many units of a perishable commodity should be purchased to meet an uncertain demand for the item, where the item is perished after two periods and the demand may be for new items or for items one period old. The states of the system are the numbers of one-period old items at the decision epoch. The new states at the next decision epoch depend upon the demands for new and one-period old items and on the decisions taken.	Expected profit per unit time over an infinite time horizon.	The author says that the problem is inspired by a doughnut advertisement offering 'one-day old doughnuts for half price'. The problem is formulated as an infinite-horizon Markov decision process, and the computations carried out using linear programming.
Seidmann and Schweitzer[55]	Decisions have to be made as to which of a set of parts a machine should make when it is free to do so, and where the parts then feed on to one of a set of further production processes. The states of the system are the numbers of completed parts waiting, or being produced, for the subsequent operations. The new states at the next decision epoch depend upon the parts completed at the second stage between the epochs, and on the decisions taken.	Expected penalties per unit time over an infinite horizon, where the penalties arise from machines being idle at the second stage.	The problem is formulated as an infinite-horizon semi-Markov stochastic dynamic program. A relative value algorithm is used to obtain optimal policies.
Burstein *et al.*[56]	Decisions have to be made as to how much of a given product to produce when the demand quantities are known, but where the timing of these demands is uncertain. The states of the system are current inventory levels and the complete histories of the demands up to the decision epoch. The new states at the next decision epoch depend upon current demands and on the decisions taken.	Expected cost over a finite horizon, where the relevant costs are production costs, inventory holding costs, and back-logging costs.	The problem is formulated as a finite-horizon stochastic dynamic program and some structural results for an optimal policy are given. The case when the demands are random is also considered.
Bartmann[57]	This problem is similar to that of Symonds[49], the essential differences being that the decisions are production decisions, and the states also include the previous production rates, and forecasts of demand for the next period.	Expected discounted cost over a finite horizon where the relevant costs are production costs (inclusive of costs for changing production rates), and inventory holding costs.	The analysis is applied to the German automotive industry with quarterly decisions. The problem is modelled as a finite-horizon stochastic dynamic program.
Golabi[58]	This problem is similar to that of Kingsman[46], the essential differences being that there is a slight change in inventory holding costs, and prices are time dependent.	As for Kingsman[46], except that, in addition, discounted costs are studied.	As for Kingsman[46], although finite and infinite horizon cases are studied. It gives conditions under which a structural policy is optimal.

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
		6. Finance and investment	
Norman and White[59]	Each day an insurance company has to make decisions as to how much of its effective bank balance it should invest, in the light of random claims, expenses, and call-off by stockbrokers. The states of the system on any day are the effective bank balances. The new states at the next decision epoch depend upon the above events and on the decisions taken.	Expected bank balance after a finite number of days. For the infinite-horizon problem the objective function is a growth rate function.	Real data are used for a UK insurance company. The problem is formulated as a finite-horizon, stochastic dynamic program. A structural optimal policy is derived. Simulation shows that actual behaviour is close to optimum.
Derman *et al.*[60]	This problem has some similarities to the problem of Saario[61], the essential difference being that any amount may be invested, rather than sold, at any time, and the returns for a specified level of investment are not subject to price variations. The states of the system at any decision epoch are the levels of holdings available for investment. The only random factor is the time interval to the next investment opportunity.	Expected profit over a finite horizon.	The problem is formulated as a finite-horizon semi-Markov stochastic dynamic program. The continuous-time case is also studied. Some structural policy results are derived.
Mendelssohn and Sobel[62]	Decisions have to be made at each decision epoch as to how much of a specified amount of capital should be invested in the light of unknown returns on investment. The states at each decision epoch are the levels of capital available. The new states at the next decision epoch depend upon the current investment returns and on the decisions taken.	Expected discounted utilities of consumption quantities over finite and infinite time horizons.	The problem is formulated both as a finite and infinite-horizon stochastic dynamic program. Under certain assumptions a structural optimal policy is derived.
Bartmann[63]	Credit institutions have to make daily decisions as to how much, and in what form, they should hold cash, in the light of possible robberies, and the need to ask for shipment of cash when shortages appear imminent.		Little information is available on the details of the problem or on the model used. Real data is used. An optimal structural policy is derived.
Wessels[64]	This is similar to the problem of Bartmann[62]. The states of the system on any day are the cash levels. The new states at the next decision epoch depend upon current deposits and withdrawals and on the decisions taken.	Expected cost per day over an infinite horizon, where the relevant costs are loss of interest costs, replenishment costs, and stock reduction costs. A constraint on cash shortages is imposed.	Real bank data are used. The problem is formulated as an infinite-horizon stochastic dynamic program. Structural policies are used.
Lee[65]	Decisions have to be made at each decision epoch, during a research and development programme, as to whether any action should be taken, and, if so, whether this should be to seek more technological information or to seek higher	Expected discounted profits over an infinite horizon, where the profits take into account the costs of information gathering.	The problem is formulated as an infinite-horizon stochastic dynamic program. Some monotone results are obtained.

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
	economic rewards with current knowledge. The states of the system at any decision epoch are indices of current technological knowledge, and of current economic knowledge. The new states at the next decision epoch depend upon changes in the latter which take place and on the decisions taken.		
Butler *et al.*[66]	A holder of an asset has to decide, in the light of offers made, whether or not to sell his asset, or to wait for a possibly better offer. The states of the system at any decision epoch are the histories of all the offers made to date. The new states at the next decision epoch depend upon any current new offer and on the decisions made.	Expected net income over a finite and infinite time horizon, where the net income takes into account a cost for delaying the decision until the next decision epoch.	The problem is formulated as a finite and as an infinite-horizon stochastic dynamic program. Two models are studied, one allowing only the latest offer to be accepted at any decision epoch, and the other allowing any of the offers made to be accepted. Structural optimal policy results are obtained.
Saario[61]	A holder of several units of some asset has to decide, at each decision epoch, whether or not to sell one unit at the current price. The states of the system at each decision epoch are, effectively, the numbers of units remaining and the current price offers. The new states at the next decision epoch depend upon any new offer, if one is made, and on the decisions taken.	Expected discounted income in disposing of all units within a specified finite-time horizon, and/or infinite horizons, where all remaining units at the deadline must be disposed of at a specified price.	The problem is formulated as a finite and as an infinite-horizon stochastic dynamic program. Structured optimal policies are derived.
Monahan and Smunt[67]	Decisions have to be made at each decision epoch as to how much the existing proportion of automated technology should be increased to meet a specified demand pattern. The states of the system at each decision epoch are the current interest rates, the current proportions of automated technology, and the levels of automated technology currently available for purchase. The new states at the next decision epoch depend upon changes in these in the current period and on the decisions taken.	Expected cost over a finite number of periods, which make allowance for a salvage value of the equipment at the end of this period.	The problem is formulated as a finite-horizon stochastic dynamic program. Special forms of cost function are studied and some structural policy results derived.
		7. *Queues*	
Low[68]	When a customer arrives or departs in a multichannel queueing system, decisions have to be made as to the price to be charged for a facility's service, which will influence the arrival rate. The states of the system at any decision epoch are the numbers in the system. The new states at the next decision	Expected reward per unit time over an infinite-horizon, where the rewards consist of customer payments and negative customer waiting costs.	The problem is formulated as an infinite-horizon semi-Markov stochastic dynamic program. A monotone result for optimal policies is derived.

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
	epoch depend upon arrival or service being the future event to occur, and on the decisions made.		
Ignall and Kolesar[69]	In a two terminal shuttle problem, with the shuttle moving between two terminals, decisions have to be made at one terminal as to when to dispatch the shuttle to the next terminal, where it never waits. The decisions are made when the shuttle arrives at the terminal or when the next passenger arrives. The states of the system at each decision epoch are the numbers of passengers at each terminal. The new states at the next decision epoch depend upon arrivals between the epochs and on the decisions made.	Expected cost per unit time over an infinite horizon, where the relevant costs are customer waiting costs and shuttle transportation costs.	The problem is formulated as an infinite-horizon semi-Markov stochastic dynamic program. It looks at three problem structures and derives structural policy results. Also mentions a reformulation of the problem to minimize waiting time subject to a constraint on shuttle frequency.
Deb[70]	This is similar to the problem of Ignall and Kolesar[69], the essential differences being that dispatch decisions are made at each terminal, the inter-terminal travel characteristics are different, the transportation costs are different, and decisions are made when a customer arrives. The states of the system also include the location of the shuttle.	Expected discounted cost over an infinite horizon, where the relevant costs are customer waiting costs and shuttle transportation costs.	The problem is formulated as an infinite-horizon stochastic dynamic program. The successive approximation method is used and some structural policy results are derived.
Mandelbaum and Yechiali[71]	A customer arriving at a queue has to decide whether to leave, to wait before joining the queue, or to join the queue. The states of the system at the next decision epoch depend upon the number of arrivals up to the next service and on the decisions made.	Expected cost until he finally leaves the system, where the relevant costs are negative reward costs paid upon completion of the service, and losses for not joining the queue.	The problem is formulated as an absorbing-state (where the process can terminate by decision) stochastic dynamic program. Structural policy results are derived.
Gonheim and Stidham[72]	Decisions have to be made as to whether or not customers arriving at each of two queues, which are in series, should be accepted into the queue. The states of the system are the numbers in each queue, the current reward for admitting a customer, and the queue at which the last arrival occurred. The new states at the next decision epoch depend on whether or not a service or arrival occurs as the next event, and on the changes in the reward per customer, as well as on the decisions made.	Expected finite and infinite-horizon discounted reward, where the relevant rewards are the customer payments and negative waiting costs.	The problem is formulated as a finite and infinite-horizon stochastic dynamic program. Some structural policy results are obtained.
Yeh[73]	At each decision epoch, decisions have to be made as	Expected finite and infinite horizon discounted costs,	The problem is formulated as a finite and infinite-horizon

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
	to which service rate should be used. The states of the system are the parameters of a Beta distribution for the probability of an arrival in the next period. The next states depend upon whether an arrival occurs or a service occurs in the next period as well as on the decisions made.	where the relevant costs are customer waiting costs and service costs.	stochastic dynamic program. Some structural policy results are obtained.
		8. *Sales promotion*	
Rao and Thomas [74]	Decisions have to be made as to the price discount, and its duration, to be offered on a product. The states of the system at each decision epoch are the numbers of discounts to date, the numbers of periods remaining on the current discount, the current prices, the amounts of each discount still remaining, and the lengths of the discount periods to date. The new states at the next decision epoch depend upon the current sales as well as on the decisions made.	Expected profit over a finite-horizon, where the profit is sales revenue net of penalty costs for exceeding a specified budget.	The problem is formulated as a finite-horizon stochastic dynamic program. No analysis is attempted.
Wessels and Van Nunen [75]	Decisions have to be made as to whether or not sales promotional plans should be initiated. The states of the system at each decision epoch are the current sales levels, and the numbers of periods since the last, and last but one, sales promotional initiatives.	Expected profit over a finite-horizon, where the profit is sales revenue net of the cost of promotional activities.	It is not clear, from the paper, as to the model used. The authors state that the paper is based upon some real experiences at Unilever.
Deshmukh and Winston [76]	A dominant firm in a market has to decide on the price of its product. The states of the system are the total sizes of the industry at the decision epoch. The new states at the next decision epoch will depend upon the increase in the size of the industry arising from the price setting.	Expected infinite-horizon discounted profit.	Continuous-time stochastic dynamic programming is used and some structural results derived.
Monahan [77]	Decisions have to be made in each of a sequence of time units as to how much should be spent on advertising a product. The states of the system are the total weighted advertising expenditures and demand to date, and the new states depend upon the current advertising expenditure.	Expected discounted profit over a finite and infinite horizon, where the profits are the sales revenues plus salvage values less advertising expenditures.	The problem is formulated as a finite and infinite-horizon stochastic dynamic program. Some structural policy results are derived.
		9. *Search*	
Ross [78]	Decisions have to be made as to which locations to search for a target. The states of the system at each decision epoch are the probabilities of the	Expected cost up until the target is found, where the relevant costs are search costs and negative reward costs.	The problem is formulated as an absorbing-state stochastic dynamic program. Structural policy results are obtained.

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
	target being in each of the locations. The new states at the next decision epoch depend upon information obtained from the search decision.		
Chew[79]	This is similar to the problem of Ross[78]. The essential difference is that the process may be terminated, at a cost, before the target is found.	Expected cost up until the target is found or until the search is terminated, where the relevant costs are search costs and termination costs.	The problem is formulated as an absorbing-state stochastic dynamic program. Structural policy results are obtained.
Eagle[80]	This problem is similar to that of Ross[78]. The essential differences are that the search paths are constrained, the objective function is different, and the states of the system also include the last location to be searched.	Probability of finding the target within a finite number of searches.	The problem is formulated as an absorbing-state, finite-horizon, stochastic dynamic program. Dominance ideas are used to eliminate non-optimal actions.
		10. *Motor insurance claims*	
Hastings[81]	When a driver is involved in a motor accident decisions have to be made as to whether or not a claim should be made. The states of the system at each decision epoch are the current category of annual premiums for the driver. The new states of the system at the next decision epoch depend upon accidents which may yet occur before the next premium renewal and on the decisions made.	Expected cost over a finite-horizon, where the relevant costs are repair costs and premium costs.	Real data based upon Royal Insurance Company policies are used. The problem is formulated as a finite-horizon, stochastic dynamic program. It is assumed that a structural optimal policy is used, but no validation of this is given. The decisions structure is equivalent to the choice of these limits.
Kolderman and Volgenant[82]	This is similar to the problem of Hastings[81]. The essential differences are that only one claim is needed to change the insurance premium category, and any extra claims in the year have no effect on the current repair estimate for an accident.	Expected cost per year over an infinite horizon, where the relevant costs are repair costs and premium costs.	Real data, based upon the 1982 Dutch Automobile insurance system, are used. The problem is formulated as an infinite-horizon stochastic dynamic program. The paper assumes that structural optimal policies are to be used, and uses policy space iteration, with no validation of either.
		11. *Overbooking*	
Rothstein[83]	Decisions have to be made each day, prior to a target day, as to whether airline bookings for that target day should be accepted. The states of the system at each decision epoch are the numbers of bookings to date for the target day. The new states at the next decision epoch depend upon acceptances and cancellations and on the decisions made.	Expected income for the target day, where the relevant income is income from realized bookings, from last-minute bookings, and negative income for not being able to honour bookings on the day. The overbooking aspect is handled by a constraint on the expected number of bookings.	Real data are used for the American Airlines, Dallas to Chicago flights. The problem is formulated as a finite horizon stochastic dynamic program. It is solved using the method of successive approximations, and a Lagrange multiplier is used to handle the overbooking constraint.
Rothstein[84]	This is similar to the problem of Rothstein[83] but applied to hotels instead of airlines.	As for Rothstein[83]	Real data are used for Sheraton Pocono Inn, Stroudsburg, Pennsylvania. The problem is formulated as a finite-horizon stochastic dynamic program. The

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
			constraint is effectively handled by Lagrange multipliers, and the method of successive approximations is used.
Ladany[85]	This is similar to the problem of Rothstein[83]. The essential difference is that consideration of bookings for single beds, and for double beds, is given. The states of the system at each decision epoch are the numbers of bookings for single beds and for double beds.	As for Rothstein[83] except that penalty costs, and not constraints, are used to handle overbookings.	As for Rothstein[83].
Gottleib and Yechiali[86]	This is similar to the problem of Rothstein[83]. The essential differences are that continuous time is used, the cancellation characteristics are different, the decisions allow acceptance and rejection of new applications, cancellations of bookings and purchasing new reservations.	Expected profit up to, and including, the target day, where profit is derived from income from realized bookings and from negative income from cancelling bookings and buying new reservations.	The problem is formulated as a finite-horizon, continuous-time, stochastic dynamic program. Optimal structural policies are derived.
Rothstein[87]			This paper includes references to stochastic programming applications to airline overbooking.
		12. *Epidemics*	
Lefevre[88]	Decisions have to be made, in the face of an epidemic, as to the quarantine level and medical treatment level which should be used. The states at each decision epoch are the numbers in the population who are infected and can transmit the disease. The new states at the next decision epoch depend upon disease propagation rates and on the decisions made.	Expected cost during the period of the epidemic, inclusive of a social cost of the epidemic.	A special transformation is used to convert the continuous time problem into a finite-state stochastic dynamic program subject to an absorbing state. Some monotone properties of an optimal policy are derived.
Jaquette[11]			This paper contains references to stochastic dynamic programming models for pest control.
		13. *Credit*	
Bierman and Hausman[89]	Decisions have to be made as to whether or not credit should be granted to a customer. The states of the system at each decision epoch are the parameters of a Beta distribution for the probability that the customer will repay the credit given. The new states of the system at the next decision epoch depend upon whether or not the customer did repay, and on the decisions made.	Expected discounted gain over a finite horizon where the gains arise from payments made by the customer and from negative gains from defaulting payments.	The problem is formulated as a finite-horizon stochastic dynamic program. The author states that the problem is, in reality, more complex than the one formulated. The method of solution is by successive approximations. An extension to decisions concerning the amount of credit to give is considered.

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
Liebman[90]	Decisions have to be made for each customer as to whether no action should be taken, or a letter requesting payment should be made. The states of the system at each decision epoch are the ages of the account, past payment history and dollar volume charges. The new states at the next decision epoch depend upon changes in these categories, and on the decisions made.	Expected discounted cost over finite and infinite-horizons, where the relevant costs are credit action costs, carrying costs, and bad debt costs.	Real data are used. The problem is formulated as a finite-horizon, and as an infinite-horizon, stochastic dynamic program. In the infinite-horizon case linear programming is used.
		14. *Sports*	
Kohler[91]	In a game of darts decisions have to made as to where next to aim for, bearing in mind errors in the actual score. The states of the system are the residual required scores to complete a player's game, and the number of his throw in a given round. The new states at the next decision epoch depend upon his score.	Expected number of throws to complete the game.	The problem is formulated as an absorbing state stochastic dynamic program. The solution method is by working backwards from zero residual score. A branch-and-bound routine is used.
Norman[92]	A tennis player has to decide whether to make a fast serve. The states of the system, in a given game, are the scores of each player, and the numbers of the services due for the serving player. The states at the next decision epoch depend upon the success or failure of the serve.	Probability of winning the game.	The problem is formulated as an absorbing state stochastic dynamic program. An optimal structural policy is derived.
		15. *Patient admissions*	
Lopez-Toledo[93]	Decisions have to be made as to whether or not a patient should be admitted to a hospital. The states of the system at each decision epoch are the hospital population sizes. The new states at the next decision epoch depend upon hospital departures and on the decisions made.	Expected discounted return over an infinite horizon.	Real data are used and policies checked via simulation.
		16. *Location*	
Rosenthal *et al.*[94]	A service facility has to be moved from one location to another to meet demands for its service. The states of the system at any decision epoch are the locations of the service facility and the locations of the customer. The new states at the next decision epoch depend upon the location of the next call for service and on the decisions taken.	Expected discounted infinite-horizon costs, where the relevant costs are relocation costs and costs of serving one location from another location.	The problem is formulated as an infinite-horizon stochastic dynamic program. The method of successive approximations is used with the Gauss Siedel variation. Consideration is given to the multi-customer case. A special heuristic is developed for this case to overcome the high cardinality of the problem.
		17. *Design of experiments*	
Kolonko and Benzing[95]	In an experimental process decisions have to be taken as	Expected discounted reward over a finite-horizon, where	The problem is formulated as a finite-horizon stochastic

TABLE 3. *Continued*

Reference	Short summary of the problem	Objective function	Comments
	to which of the two experiments to use. The states of the system at each decision epoch are the numbers of the experiment last used, and the numbers of successes and failures for the second experiment up to the decision epoch. The new states at the next decision epoch depend upon whether the second experiment is a failure or a success and on the decisions taken.	the rewards are from the experiments plus the negative rewards from switching from one experiment to the other.	dynamic program. Some structural policy results are derived.
		18. *General*	
Ellis *et al.*[96]			This paper contains references to material on applications of stochastic dynamic programming.
White[1,2]			As for Ellis *et al.*[96]
Jarrow and Rudd[97]			This book contains material on both discrete and continuous-time option pricing, which is related to Markov decision processes, but treated in a different manner to those of the other references of this paper.
Rubenstein and Cox[98]			As for Jarrow and Rudd[97].

COMMENTS ON THE SURVEY

We will deal with these comments under several headings.

Real life data

The breakdown of papers which make explicit reference to the use of real data, although not to any implementation, is given in Table 4.

TABLE 4. *Real life data*

1 *Population harvesting*: Mendelssohn[4-6], Ben-Ari and Gal[8].
2 *Agriculture*: Brown *et al.*[9], Onstad and Rabbinge[10].
3 *Water resources*: Little[14], Arunkumar and Chon[18], Turgeon[20-22], Krzysztofowicz and Jaganathan[23], Krzysztofowicz[28].
4 *Inspection, maintenance and repair*: None.
5 *Purchasing, inventory and production*: Bartmann[57].
6 *Finance and investment*: Norman and White[59], Bartmann[62], Wessels[75].
7 *Queueing*: None.
8 *Sales promotion*: None.
9 *Search*: None.
10 *Motor insurance claims*: Hastings[81], Kolderman and Volgenant[82].
11 *Overbooking*: Rothstein[83,84].
12 *Epidemics*: None.
13 *Credit*: Liebman[90].
14 *Sports*: None.
15 *Patient admissions*: Lopez-Toledo[93].
16 *Location*: None.
17 *Design of experiments*: None.
18 *General*: Ellis *et al.*[96], *White*[1,2].

Since the sample is clearly unrepresentative of the true situations, no conclusions can be drawn. It is interesting to note the significantly greater concern with real data, as a proportion of papers in each section, in the areas of population harvesting, agriculture, water resources and motor insurance, as against those in inspection, maintenance and repair and in queueing.

Structural Results

Those papers which give structural results, in one of several forms, are given in Table 5.

TABLE 5. *Structural results*

1	*Population harvesting*: Mendelssohn[4,5], Mann[7].
2	*Agriculture*: None.
3	*Water resources*: Russell[17], Arunkumar and Chan[18].
4	*Inspection, maintenance and repair*: Gluss[29], White[30], Duncan and Scholnick[36], Crabhill[37], Butler[38], Stengos and Thomas[39], Sengupta[40], Hayre[41], Tijms[42], Ohnishi *et al.*[43].
5	*Purchasing, inventory and production*: Kingsman[46], Sobel[47], Kalymon[48], Thomas[51], Karmarkar[53], Burstein *et al.*[56], Golabi[58].
6	*Finance and investment*: Norman and White[59], Derman *et al.*[60], Mendelssohn and Sobel[61], Bartmann[62], Wessels[63], Lee[64].
7	*Queueing*: Low[68], Ignall and Kolesar[69], Deb[70], Mandelbaum and Yechiali[71], Gonheim and Stidham[72], Yeh[73].
8	*Sales promotion*: Deshmukh and Winston[76], Monahan[77].
9	*Search*: Ross[78], Chew[79].
10	*Motor insurance claims*: Hastings[81], Kolderman and Volgenant[82].
11	*Overbooking*: Gottleib and Yechiali[86].
12	*Epidemics*: Lefevre[88].
13	*Credit*: None.
14	*Sports*: Norman[92].
15	*Patient admissions*: None.
16	*Location*: None.
17	*Design of experiments*: Kolonko and Benzing[95].
18	*General*: Several references in the cited articles contain structural results.

It is not intended, in this paper, to describe the voluminous structural result material in the many papers cited. These results can take many forms. For example (see Mendelssohn[6]), if x_n is the vector of population numbers, each component corresponding to a particular category of the population, there exists a base stock vector x_n^0, where n is the number of periods remaining, such that the optimal policy is to leave y_n to breed where

$$y_n = \min[x_n, x_n^0]$$

where 'min' means 'component minimization.'

In Kolonko[75] the result is that, once experiment number one has been used, it is always used.

In Norman[92] it is possible to give explicitly the probability regions within which specific policies are optimal.

Deshmukh and Winston[76] show that, after a certain time, the optimal price is monotone non-decreasing in time.

In Lee[64] it is shown that investment in research and development will be monotone non-decreasing in the discount factor.

These are simply illustrations of a vast variety of results. These results are valuable in the following manner namely:

(i) they facilitate the computation of optimal policies;
(ii) even though the real-life problems which relate to many of the actual models used are more complex, they can be used to provide insight into the selection of policies which might be considered for the more complex problems, and which might be solved by other means, e.g. by simulation. They also provide a very rich ground for mathematical studies in this area.

Some of the papers provide the basis for interesting research questions in this context, e.g.

(i) in some papers, specific structural policies are used, but the papers do not demonstrate that optimal policies can be found within the specified class of policies, namely, Arunkumar and

Chon[18], Tijms and van der Duyn Schouten[42], Thomas[51], Wessels[63], Hastings[81], Kolderman and Volgenant[82];

(ii) in the cases of Tijms and van der Duyn Schouten[42] and Kolderman and Volgenant[82], the structured policies are used in the context of restricted policy space iteration, although, in general, policy iteration does not preserve structure.

Both (i) and (ii) pose research questions; in the former case the problem is to validate the structures used, and in the second case, the problem is to validate that, if a structural optimal policy exists, restricted policy space will find it under some circumstances.

Computational aspects

Very few of the papers pose serious computational problems. However, a few do and sometimes special approaches are developed to tackle these, which may have a more general merit. The relevant papers are given in Table 6.

TABLE 6. *Special computation schemes*

Reference	Scheme
Mendelssohn[4,5]	Decomposition of the state-space is possible because of the special structure of the problem. The method is exact.
Ben-Ari and Gal[8]	Functional approximation is used to overcome the high state-space cardinality difficulty.
Turgeon[20]	A special approximating decomposition is used to overcome the high state-space cardinality difficulty.
Turgeon[21]	A special approximating aggregation procedure is used to overcome the high state-space cardinality difficulty.
Turgeon[22]	An aggregation of the state-space is possible because of the special structure of the problem. The method is exact.
Sobel[47]	In this problem a design parameter has to be chosen and a special parametric scheme is developed.
White[50]	Weak coupling of successive weeks allows an approximating algorithm to be developed for a problem with a high cardinality state-space.
Rosenthal *et al.*[94]	A heuristic decomposition method is developed to overcome the high state-space cardinality difficulty. It is conjectured that the effectiveness of the heuristic increases as the cardinality increases.

Although the number of references cited here is small, it is clear that some headway is being made to overcome the difficulties posed by high-cardinality state-spaces. In particular, the functional approximation method of Ben-Ari and Gal[8], which was proposed by Bellman in the 1950s, is, perhaps, a method which might be given further attention.

Problem realism

There are clearly many ways in which more realism might be incorporated into the models, and this is explicitly recognized by some authors and, almost certainly, implicitly by all the authors. It is not the purpose of this paper to deal with these, since every paper would be subject to more realistic modifications. There are three general ways in which more realism might be incorporated into some of the papers. These are as follows, and give rise to research questions.

Infinite horizon models. It is quite unrealistic to use infinite-horizon models without some justification. However, the policies obtained from infinite horizon models may be approximately optimal for some finite-horizon situations and the research question is to find ways of determining how good an approximation they are, or, if they are used as a starting policy in some iterative procedure to solve the finite-horizon problem, then how good would the procedure be?

Note that this is the inverse of the usual procedure of approximating infinite horizon processes by finite horizon schemes.

Measure of performance. In all cases, the objective functions are the expected discounted cost, or reward, or the expected cost, or reward, per unit time, with the addition, in some cases, of

specified constraints on certain aspects of behaviour of the system. In a few cases utility functions are used.

In almost all cases it would be difficult to justify the criteria used as a basis for supportable action, and some attempts can be made to introduce more realistic criteria, either by using appropriate utility functions, by adding appropriate constraints, or by adding variance considerations. This would invariably make the computational issues more complex, but it might increase the rate of implementation of such modelling.

The Markov assumption. In every case, one might justifiably question whether, in terms of the state structure used, the state transitions are Markov. Indeed, it is possible that this aspect constitutes one of the greater barriers to implementation. On the other hand, involving more history in the state description in order to try to achieve a Markov process, may make the model computationally intractable. This raises the research question of how we can determine, for a given problem, how good an approximation an appropriate Markov assumption is for the original state space.

REFERENCES

1. D. J. White (1985) Real applications of Markov decision processes. *Interfaces* **15**(6), 73-78.
2. D. J. White (1988) Further real applications of Markov decision processes. *Interfaces* **18**(5), 55-61.
3. S. Dreyfus (1957) A note on an industrial replacement process. *Opl Res. Q.* **8**, 190-193.
4. R. Mendelssohn (1978) Optimal harvesting strategies for stochastic single-species, multi-stage class models. *Mathe. Biosci.* **41**, 159-174.
5. R. Mendelssohn (1980) Managing stochastic multi-species models. *Mathe. Biosci.* **49**, 249-261.
6. R. Mendelssohn (1982) Discount factors, risk aversion in managing random fish populations. *Canadian J. Fisheries and Aquatic Sci.* **39**, 1252-1257.
7. S. H. Mann (1970) A mathematical theory for the harvest of natural animal populations when birth rates are dependent upon total population size. *Mathe. Biosci.* **7**, 97-110.
8. Y. Ben-Ari and Shmuel Gal (1986) Optimal replacement policy for multicomponent systems: an application to a dairy herd. *Eur. J. Opl Res.* **23**, 213-221.
9. B. G. Brown, R. W. Katz and A. H. Murphy (1984) Case studies of the economic value of monthly and seasonal climate forecasts, 1: the planting/fallowing problem. Department of Atmospheric Sciences, Oregon State University, Corvallis, Oregon, USA.
10. D. W. Onstad and R. Rabbinge (1985) Dynamic programming and the computation of economic injury levels for crop disease control. *Ag. Syst.* **18**, 207-226.
11. D. L. Jaquette (1972) Mathematical models for controlling growing biological populations: a survey. *Opns Res.* **20**, 1142-1151.
12. G. R. Conway (1973) Experience in insect pest modelling: a review of models, uses and future directions. In *Insects: Studies and Population Management.* (P. W. Geir, L. R. Clark, D. J. Anderson, H. A. Nix, Eds) pp 103-130. Ecological Society of Australia, Canberra.
13. R. M. Feldman and G. L. Curry (1982) Operations research for agricultural pest management. *Opns Res.* **30**, 601-618.
14. J. D. C. Little (1955) The use of storage water in hydroelectric systems. *Opns Res.* **3**, 187-197.
15. N. Buras (1965) A three dimensional optimisation problem in water-resources engineering. *Opl Res. Q.* **16**, 419-428.
16. S. Y. Su and R. A. Deininger (1972) Generalisation of White's method of successive approximation to periodic Markovian decision processes. *Opns Res.* **19**, 318-326.
17. C. B. Russell (1972) An optimal policy for operating a multi-purpose reservoir. *Opns Res.* **20**, 1181-1189.
18. S. Arunkumar and K. Chon (1978) On optimal regulation policies for certain multi-reservoir systems. *Opns Res.* **26**, 551-562.
19. U. Shamir (1980) Application of operations research in Israel's water sector. *Eur. J. Opl Res.* **5**, 332-345.
20. A. Turgeon (1980) Optimal operation of multireservoir power systems with stochastic inflows. *Water Resources Res.* **16**, 275-283.
21. A. Turgeon (1981) A decomposition method for the long-term scheduling of reservoirs in series. *Water Resources Res.* **17**, 1565-1570.
22. A. Turgeon (1985) The weekly scheduling of hydroelectric power plants subject to reliability constraints. Hydro-Quebec, Canada.
23. R. Krzysztofowicz and E. V. Jagannathan (1981) Stochastic reservoir control with multiattribute utility criterion. *Proc. International Symposium on Real-Time Operation of Hydrosystems*, University of Waterloo, Canada, 145-159.
24. S. Yakowitz (1982) Dynamic programming applications in water resources. *Water Resources Res.* **18**, 673-696.
25. W. G. Yeh (1983) Optimisation models for reservoir operation: a state-of-the-art review. School of Engineering and Applied Science, UCLA, California.
26. S. C. Sarin and W. El Benni (1982) Determination of optimal pumping policy of municipal water plant. *Interfaces.* **12**(2), 43-48.
27. J. R. Stedinger, B. F. Sule and D. P. Loucks (1984) Stochastic dynamic programming models for reservoir operation optimisation. *Water Resources Res.* **20**, 1499-1505.
28. R. Krzysztofowicz (1986) Optimal water supply planning based on seasonal runoff forecasts. *Water Resources Res.* **22**, 313-321.

29. B. Gluss (1959) An optimal policy for detecting a fault in a complex system. *Opns Res.* **24**, 468-477.
30. D. J. White (1962) Optimal revision periods. *J. Mathe. Anal. and Applications* **4**, 353-365.
31. D. J. White (1965) Dynamic programming and systems of uncertain duration. *Mgmt Sci.* **19**, 37-67.
32. D. J. White (1967) Setting maintenance inspection intervals using dynamic programming. *J. Ind. Eng.* **XVIII**, 376-381.
33. J. E. Eckles (1968) Optimum maintenance with incomplete information. *Opns Res.* **16**, 1058-1067.
34. H. Hinomoto (1971) Sequential control of homogeneous activities—linear programming of semi Markovian decisions. *Opns Res.* **19**, 1664-1674.
35. E. P. Kao (1973) Optimal replacement rules when changes of state are semi-Markovian. *Opns Res.* **21**, 1231-1249.
36. J. Duncan and L. S. Scholnick (1973) Interrupt and opportunistic replacement strategies for systems of deteriorating components. *Opl Res. Q.* **24**, 271-283.
37. T. B. Crabhill (1974) Optimal control of a maintenance system with variable service rates. *Opns Res.* **22**, 736-745.
38. D. A. Butler (1979) A hazardous inspection model. *Mgmt Sci.* **25**, 79-89.
39. D. Stengos and L. C. Thomas (1980) The blast furnaces problem. *Eur. J. Opl Res.* **4**, 330-336.
40. B. Sengupta (1981) Control limit policies for early detection failure. *Eur. J. Opl Res.* **7**, 257-264.
41. L. S. Hayre (1983) A note on optimal replacement policies for deciding whether to repair or replace. *Eur. J. Opl Res.* **12**, 171-175.
42. H. C. Tijms and F. A. Van der Duyn Schouten (1984) A Markovian algorithm for optimal inspections and revisions in a maintenance system with partial information. *Eur. J. Opl Res.* **21**, 245-253.
43. M. Ohnishi, H. Kawai and H. Mine (1986) An optimal inspection and replacement policy under incomplete state information. *Eur J. Opl Res.* **27**, 117-128.
44. L. C. Thomas (1986) A survey of maintenance and replacement models for maintainability and reliability of multi-item systems. *Reliability Eng.* **16**, 297-309.
45. D. P. Gaver, P. Jacobs and L. C. Thomas (1987) Optimal inspection policies for standby systems. *Communications in Statistics—Stochastic Models* **3**, 259-273.
46. B. G. Kingsman (1969) Commodity purchasing. *Opl Res. Q.* **20**, 59-79.
47. M. J. Sobel (1970) Making short-run changes in production when the employment level is fixed. *Opns Res.* **18**, 35-51.
48. B. A. Kalymon (1971) Stochastic prices in a single-item inventory purchasing model. *Opns Res.* **19**, 1434-1458.
49. G. H. Symonds (1971) Solution method for a class of stochastic scheduling problems for the production of a single commodity. *Opns Res.* **19**, 1459-1466.
50. D. J. White (1973) An example of loosely coupled stages in dynamic programming. *Mgmt Sci.* **19**, 739-746.
51. L. J. Thomas (1974) Price production decisions with random demand. *Opns Res.* **22**, 513-518.
52. R. D. Snyder (1975) A dynamic programming formulation for continuous time stock control systems. *Opns Res.* **23**, 383-386.
53. U. S. Karmarkar (1981) The multiperiod multilocation inventory problem. *Opns Res.* **29**, 215-228.
54. M. Parlar (1982) Optimal policies for a perishable and substitutable product: a Markov decision model. Faculty of Administration, University of New Brunswick.
55. A. Seidmann and P. J. Schweitzer (1982) Part selection policy for a flexible manufacturing cell feeding several production lines. Working Paper, QM 8217, Graduate School of Management, University of Rochester.
56. M. C. Burstein, C. H. Nevison and R. C. Carlson (1984) Dynamic lot-sizing when demand timing is uncertain. *Opns Res.* **32**, 362-379.
57. D. Bartmann (1984) Mittelfristige productionsplanung bei ungewissen szenarieb. *Z. Opns Res.* **28**, B187-B204.
58. K. Golabi (1985) Optimal inventory policies when ordering prices are random. *Opns Res.* **33**, 575-588.
59. J. M. Norman and D. J. White (1965) Control of cash reserves. *Opl Res. Q.* **16**, 309-328.
60. C. Derman, G. J. Lieberman and S. M. Ross (1984) A stochastic sequential allocation model. *Opns Res.* **32**, 362-379.
61. R. Mendelssohn and M. Sobel (1980) Capital accumulation and the optimisation of renewable models. *J. Econ. Theory* **23**, 243-260.
62. D. Bartmann (1980) The optimal regulation of the cash-balances of a credit institution. *Proc. Opns Res.* **9**, 77-78.
63. J. Wessels (1980) Markov decision processes; implementation aspects. Memorandum COSOR 80-14, Department of Mathematics, Eindhoven.
64. T. K. Lee (1982) On the reswitching property of R&D. *Mgmt Sci.* **28**, 887-899.
65. D. A. Butler, R. D. Shapiro and D. B. Rosenfield (1983) Optimal strategies for selling an asset. *Mgmt Sci.* **29**, 1051-1061.
66. V. Saario (1985) Limiting properties of the discounted house-selling problem. *Eur. J. Opl Res.* **20**, 206-210.
67. G. E. Monahan and T. L. Smunt (1985) A Markov decision process model of the automated flexible manufacturing investment decision. Graduate School of Business Administration, Washington University, St. Louis, Missouri 63130.
68. D. W. Low (1974) Optimal dynamic pricing policies for an *M/M/S* Queue. *Opns Res.* **22**, 545-561.
69. E. Ignall and P. Kolesar (1974) Optimal dispatching of an infinite-capacity shuttle: control at a single terminal. *Opns Res.* **22**, 1008-1024.
70. R. K. Deb (1978) Optimal dispatching of a finite capacity shuttle. *Mgmt Sci.* **24**, 1362-1372.
71. A. Mandelbaum and U. Yechiali (1983) Optimal entering rules for a customer with wait option at an *M/G/*1 queue. *Mgmt Sci.* **29**, 174-187.
72. H. A. Gonheim and S. Stidham (1985) Control of arrivals to two queues in series. *Eur. J. Opl Res.* **21**, 399-409.
73. L. Yeh (1985) Dynamic programming applications in water resources. *Water Resources Res.* **18**, 673-696.
74. V. R. Rao and L. J. Thomas (1973) Dynamic models for sales promotion policies. *Opl Res. Q.* **24**, 403-417.
75. J. Wessels and J. A. E. E. Van Nunen (1973) Dynamic planning of sales promotion by Markov programming. *Proc. XX International Meeting, The Institute of Management Sciences*, Tel Aviv, 737-742.
76. S. D. Deshmukh and W. Winston (1979) Stochastic control of competition through prices. *Opns Res.* **27**, 583-594.
77. G. E. Monahan (1983) Optimal advertising with stochastic demand. *Mgmt Sci.* **29**, 106-117.
78. S. M. Ross (1969) Problem in optimal search and stop. *Opns Res.* **17**, 984-992.
79. M. C. Chew (1973) Optimal stopping in a discrete search problem. *Opns Res.* **21**, 741-747.

80. J. N. Eagle (1984) The optimal search for a moving target when the search path is constrained. *Opns Res.* **32**, 1107–1115.
81. N. A. J. Hastings (1976) Optimal claiming on vehicle insurance. *Opl Res. Q.* **27**, 908-913.
82. J. Kolderman and A. Volgenant (1985) Optimal claiming in an automobile insurance system with Bonus-Malus structure, *J. Opl Res. Soc.* **36**, 239-247.
83. M. Rothstein (1971) An airline overbooking model. *Transp. Sci.* **5**, 180–192.
84. M. Rothstein (1974) Hotel overbooking as a Markovian sequential decision process. *Decis. Sci.* **5**, 389–404.
85. S. P. Ladany (1976) Dynamic operating rules for motel reservations. *Decis. Sci.* **7**, 829–840.
86. G. Gottleib and U. Yechiali (1983) The hotel overbooking problem. RAIRO *Recherche Operationnelle/Operations Research* **17**, 343-355.
87. M. Rothstein (1985) OR and the airline overbooking problem. *Opns Res.* **33**, 237–248.
88. C. Lefevre (1981) Optimal control of a birth and death epidemic process. *Opns Res.* **29**, 971-982.
89. H. Bierman Jr and W. H. Hausman (1970) The credit granting decision. *Mgmt Sci.* **16**, B 519–B 532.
90. L. H. Liebman (1972) A Markov decision model for selecting credit control policies. *Mgmt. Sci.* **18**, B 519-B 525.
91. D. Kohler (1982) Optimal strategies for the game of darts. *J. Opl Res. Soc.* **33**, 871–884.
92. J. M. Norman (1985) Dynamic programming in tennis—when to use a fast serve. *J. Opl Res. Soc.* **36**, 75–77.
93. A. A. Lopez-Toledo (1976) A controlled Markov chain model for nursing homes. *Simulation* **27**, 161-169.
94. R. E. Rosenthal, J. A. White and D. Young (1978) Stochastic dynamic location analysis. *Mgmt Sci.* **24**, 645-653.
95. M. Kolonko and H. Benzing (1985) The sequential design of Bernoulli experiments including switching costs. *Opns Res.* **33**, 412–426.
96. C. Ellis, D. Lethbridge and A. Ulph (1974) The application of dynamic programming in United Kingdom companies. *Omega* **2**, 533-541.
97. R. A. Jarrow and A. Rudd (1983) *Options Pricing*. Irwin, Homewood, Illinois.
98. M. Rubenstein and J. Cox (1980) *Options Markets*. Prentice-Hall, Englewood Cliffs, NJ.