Simulation Modeling In Markovian Decision Theory: A Case Study of The Gardener's Problem

Dr. Eme L.C. (B.Sc, PGD, PGDE, M.Eng. Ph.D), Paul P. Akpan (B.Sc)
Civil Engineering Department,
Anambra State University,
Uli-Nigeria.
Leeworks2002@yahoo.com

Dr. Uju I. U. (B.Sc, M.Sc, Ph.D)
Works Department,
Anambra State University,
Uli-Nigeria.

Abstract. This paper aims at studying simulation modeling in Markovian Decision theory considers its relationship to linear programming and adapts exhaustive enumeration method, policy iteration methods of certain stochastic systems of the finite and infinite stage models for solution of the gardener's problems. The objective of the problem is to determine the optimal policy or strategy or action that maximizes the expected return (revenue) within the available limited fund over the planning period. Consequently, most of the problems are decision problems for the decision maker (the gardener) such as: (as" apply fertilizer or do not apply fertilizer" (b)"whether the gardening activity will continue for a limited number of years or indefinitely". In the basic concept of Markovian Decision theory, the number of transitional probabilities and computational efforts required to solve a Markov chain grows exponentially with the number of states. The linear programming formulation in this paper is interesting, but it not as efficient computationally as the exhaustive enumeration method or the policy iteration algorithm methods of markovian decision problems, particularly for large values of stationary policies. In conclusion ,alternatively contingency and reliability tests were performed as a check which show no significant difference between the experimental and theoretical expected results and led to the acceptance of null hypothesis.

Key words: Probabilities, Markov-chain, Exponential growth, experimentation, theoretical hypothesis.

1.INTRODUCTION

In the definition of problem, according to Taha (2002),and Eme (2004) most of the problems are decision problems for the decision maker (the gardener) regarding:

(a) the option of using or not using fertilizer, again or loss depends on the decision made.

(b) whether the gardening activity will continue for a limited number of years or indefinitely. These situations are referred to as finite-stage and infinite stage decision problems.

(c) Evaluating the expected revenue resulting from pre-specified course of action for a given state of the system. For example fertilizer may be applied when ever the soil condition is poor (state 3). The decision – making process in this case is said to be represented by a stationary policy.

2. DISCUSSION OF CONCEPT OF MARKOVIAN DECISION THEORY

In developing this concept, effort is concentrated on the use of the Gardner example, due to the nature of life problems encountered in most situation of conflict and to make it better understandable.

Markovian decision theory according to Eme (2004) was developed by Andrei Markov in [1856 – 1922]. He was a distinguished Russian mathematician. who developed the new mathematical tools for inventory modeling of all statistical probability. And the problems resulting from the game theory and dynamic programming motivated the improvement of the discipline such as Markovian decision theory and provides a new basis of inventory theory. Markov process originated in the problem formulated by Francis Galton.

The use of the gardener example throughout this paper with the underlying philosophy that the example paraphrases several important applications in the areas of real life, inventory, maintenance, replacement, cash flow management, and regulation of electric power, hydro and water resources engineering.

This paper presents an application of dynamic programming to the solution of a stochastic decision process that can be described by a finite or infinite number of states. The transition probabilities between the states are described by a Markov chain. The reward structure of the process is also describe by a matrix whose individual elements represents the revenue (or cost) resulting from moving one state to another both the transition and revenue matrices depend on the decision alternatives available to be the decision make. And in the multi- purpose / and multi- objective nature of this paper, the purposes and the objective are in confliction to be satisfied with available limited resources. Therefore, the objective of the problem is to determine the optimal policy or strategy, or action that maximizes the expected revenue over a finite or infinite number of stages.

Every year at the beginning of a season, a gardener applies chemical tests to check soil condition. Depending on the outcome of the outcome of the tests the gardener productivity for the new season falls in one of three states: (1) good (2) fair and (3) poor.

Over the years the Gardner observed that current year's productivity depends only on last year's soil condition. The transition probabilities over a-1-year period from one productivity state to another can be represented in terms of the following Markov chain

$$
\begin{array}{ll}
 & \overbrace{\quad\begin{array}{ccc} 1 & 2 & 3 \end{array}\quad}^{\text{State of the System next year}} \\
\text{State of the system this year} \left\{ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \right. & \begin{pmatrix} .2 & .5 & .3 \\ 0 & .5 & .3 \\ 0 & 0 & 1 \end{pmatrix} = P^1
\end{array}
$$

.The transition probabilities in P_1 indicate that the productivity for a current year can be not better than last year's, for example, if the soil condition for this year is fair (state 2), next year's productivity. 5, or become poor (state 3), also with probability. 5.

The gardener can alter the transition probabilities p_1 by invoking other courses of action. Typically fertilizer is applied to boost the soil condition, which yields the following transition matrix p^2

$$
P^2 = \begin{array}{c|c}
 & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} .2 & .5 & .3 \\ 0 & .5 & .3 \\ 0 & 0 & 1 \end{pmatrix}
\end{array}
$$

To put the decision problem in perspective, the gardener associates a return function (or a reward structure) with the transition from one state to another. The return function expresses the gain or loss during a 1-year period, depending on the states between which the transition is made. Because the gardener has the option of using or not using fertilizer, gain and loses vary depending on the decision made. The matrices R^1 and R^2 summarize the return functions in millions of naira associated with the matrices P^1 and P^2, respectively

$$
R^1 = \left\| r^{1}{}_{i}^{j} \right\| = \begin{array}{c|c}
 & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} .2 & .5 & .3 \\ 0 & .5 & .3 \\ 0 & 0 & 1 \end{pmatrix}
\end{array}
$$

$$
\begin{array}{c|c}
 & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\
1 & \begin{pmatrix} .2 & .5 & .3 \\ & & \\ & & \end{pmatrix}
\end{array}
$$

$$R^2 = \left\| r^{2}{}_{i}^{j} \right\| = \begin{matrix} 2 & 0 & .5 & .3 \\ 3 & 0 & 0 & 1 \end{matrix}$$

The elements $r^{2}{}_{ij}{}^{1}$ of R^2 consider the cost of applying the fertilizer. For example, if the system is in state I and remains in state I during next year, its gain will be $r_{1j}{}^{2} = 6$ compared with $r^{1}{}_{11} = 7$ when no fertilizer is applied. Therefore, what kind of a decision problem does the gardener have?

First, we must know whether the gardening activity will continue for a limited number of years or indefinitely. These situations are referred to as finite and infinite – stage decision problems. In both cases, the gardener would determine the best course of action (fertilize or do not fertilize) given the outcome of the chemical tests (state of the system). The optimization will be based on the maximization of expected revenue.

The gardener may also be interested in evaluating the expected revenue resulting from pre-specified course of action for a given state of the system. For example, fertilizer may be applied whenever the soil condition is poor (stage 3). The decision making process in this case is said to be represented by a stationary policy.

Each stationary policy will be associated with a different transition and return matrices, which are constructed from the matrices P^1, P^2, R^1 and R^2

For example, for the stationary policy calling for applying fertilizer only when the soil condition is poor (state 3), the resulting transition and return matrices are given as

$$P^1 = \begin{pmatrix} .2 & .5 & .3 \\ 0 & .5 & .5 \\ 0 & 0 & .1 \end{pmatrix}, \qquad R^1 \begin{pmatrix} .7 & 6 & 3 \\ 0 & 5 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

These matrices differ from P^1 and R^2 in the third rows only, which are taken directly from P^2 and R^2 . The reason is that P^2 and R^2 are the matrices that result when fertilizer is applied in every state.

2.1 Finite-Stage Model

Suppose that the gardener plans to "retire" from gardening in N years. This model is interested in determining the optimal course of action for each year (to fertilize or not to fertilize). Optimality here calls for accumulating the highest expected revenue at the end of N years.

Let K= 1 and 2 represent the two courses of action (alternatives) available to the gardener. The matrices P^k and R^k representing the transition probabilities and reward function for alternative K are given in pages 130 and 131, are summarized here for convenience.

$$P^1 = \left\| P^1_{ij} \right\| = \begin{pmatrix} .2 & .5 & .3 \\ 0 & .5 & .5 \\ 0 & 0 & 1 \end{pmatrix}, R^1 = \left\| r^1_{ij} \right\| = \begin{pmatrix} .7 & 6 & 3 \\ 0 & 5 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

$$P^2 = \left\| P^2_{ij} \right\| = \begin{pmatrix} .3 & .6 & .1 \\ .1 & .6 & .3 \\ .05 & .4 & .55 \end{pmatrix} \quad R^2 = \left\| r^2\, i^j \right\| = \begin{pmatrix} 6 & 5 & -1 \\ 7 & 4 & 0 \\ 6 & 3 & -2 \end{pmatrix}$$

The gardener's problem is expressed as a finite-stage dynamic programming (DP) model as follows. For the sake of generalization. Suppose that the number of states for each stage (year) is m (=3 in the gardener's example) and define

F_n (i) = optimal expected revenue of stages n, n + 1, …, N, given that the state of the system (soil condition) at the beginning of year n is i

The backward recursive equation relating f_n and f_{n+1} can be written as

$$fn(i) = \max_k \left\{ \sum_{j=1}^{m} p^k_{ij} \left[r^k_{ij} + f_{n+1}(j)\right] \right\}, n= 1,2,\ldots, N. \qquad \ldots 3.1.1$$

Where f_{N+1} (j) = 0 for all j.

A justification for the equation is that the cumulative revenue, $r^k_{ij} + f_{n+1}$ (j),

resulting from reaching state j at stage n+1 from state i at stage n occurs with probability p^k_{ij} letting.

$$v^k_i = \sum_{j=1}^{m} p^k_{ij}\, r^k_{ij}$$

The DP recursive equation can be written as

$$fn(i) = \max_k \{ V^k_j \}$$

$$fn(i) = \max \left\{ v^k_j + \sum^{m} p^k_{ij}\, f_{n+1}(j) \right\}, n= 1,2,\ldots, N-1. \qquad \ldots 3.1.2$$

k j=1

2.2 The Gardener's Problem Case 1

In this case, we solve the gardener's problem using the data summarized in the matrices P^1, P^2, R^1, and R^2, given a horizon of 3 year only (N=3). Because the values of v^k_i will be used repeatedly in the computations, they are summarized here for convenience.

Table 2.1

Recall that k=1 represents "do not fertilize" and k=2 represent "fertilize"

i	v^1_i	v^2_i
1	5.3	4.7
2	3	3.1
3	-1	.4

2. 3 Infinite Stage Model

The steady-state behavior of a Markovian process is independent of the initial state of the system. This model is interested in evaluating policies for which the associated Markov chains allow the existence of a steady-state

Solution. Section 3.6 provides the conditions under which a Markov chain can yield steady-state probabilities.There are two methods for solving the infinite-stage problem. The first method calls for evaluating all possible stationary policies of the decision problem. This is equivalent to an exhaustive enumeration process and can be used only if the number of stationary policies is reasonably small.

The second method, called policy iteration, is generally more efficient, also a validity and reliability test for the first method, this is when policy iteration method is without and with discounting respectively, because it determines the optimum policy iteratively.

2. 4 Exhaustive Enumeration Method

Suppose that the decision problem has total of S stationary policies, and assume that $\mathbf{p}^s$, $\mathbf{R}^s$ are the (one-step) transition and revenue matrices associated with the policy, s =1,2,….., s. The steps of the enumeration method are as follows.

Step 1. Compute v^s_i, the expected one-step (one-period) revenue of policy s given state i, i= 1,2,…., m

Step 2. Compute π^s_i, the long-run stationary probabilities of the transition matrix $\mathbf{p}^s$ association with policy's. These probabilities, when they exist, are computed from the equations.

$$\pi^s p^s = \pi^s \qquad \text{... 3.3.1}$$

$$\pi^s_1 + \pi^s_2 + \ldots\ldots + \pi^s_m = 1 \qquad \ldots 3.3.2$$

where $\pi^s = (\pi^s_1, \ \pi^s_2, \ldots\ldots. \ \pi^s_m)$.

Step 3. Determine E^s, the expected revenue of policy s per transition step

(period), by using the formula.

$$E^s = \sum_{i=1}^{m} \pi^s_1 \, v^s_i$$

Step 4. the optimal policy s* is determine such that

$$E^{s^*} = \max_{S} \{E^s\}$$

We illustrate of the method by solving the gardener problem for an infinite period planning horizon.

2. 5 The Gardener's Problem Case II

Table 2.2. The gardener's problem has a total of eight stationary policies, as the following table shows:

Stationary Policy, s	Action
1	Do not fertilize at all
2	Fertilize regardless of the state
3	Fertilize whenever in state 1
4	Fertilize whenever in state 2
5	Fertilize whenever in state 3
6	Fertilize whenever in state 1 or 2
7	Fertilize whenever in state 1 or 3
8	Fertilize whenever in state 2 or 3

The matrices P^S and R^S for polices 3 through 8 are derived from these of policies 1 and 2. we thus have

$$P^1 = \begin{pmatrix} .2 & .5 & .3 \\ 0 & .5 & .5 \\ 0 & 0 & 1 \end{pmatrix}, R^1 = \begin{pmatrix} 7 & 6 & 3 \\ 0 & 5 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

$$P^2 = \begin{pmatrix} .3 & .6 & .1 \\ .1 & .6 & .3 \\ .05 & .4 & .55 \end{pmatrix}, R^2 = \begin{pmatrix} 6 & 5 & -1 \\ 7 & 4 & 0 \\ 6 & 3 & -2 \end{pmatrix}$$

$$P^3 = \begin{pmatrix} .3 & .6 & .1 \\ 0 & .5 & .5 \\ 0 & 0 & 1 \end{pmatrix}, R^3 = \begin{pmatrix} 6 & 5 & -1 \\ 0 & 5 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

$$P^4 = \begin{pmatrix} .2 & .5 & .3 \\ .1 & .6 & .3 \\ 0 & 0 & 1 \end{pmatrix}, R^4 = \begin{pmatrix} 7 & 6 & 3 \\ 7 & 4 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$P^5 = \begin{pmatrix} .2 & .5 & .3 \\ 0 & .5 & .5 \\ .05 & .4 & .55 \end{pmatrix}, R^5 = \begin{pmatrix} 7 & 6 & 3 \\ 0 & 5 & 1 \\ 6 & 3 & -2 \end{pmatrix}$$

$$P^6 = \begin{pmatrix} .3 & .6 & .1 \\ .1 & .6 & .3 \\ 0 & 0 & 1 \end{pmatrix}, R^6 = \begin{pmatrix} 6 & 5 & -1 \\ 7 & 4 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$P^7 = \begin{pmatrix} .3 & .6 & .1 \\ 0 & .5 & .5 \\ .05 & .4 & .55 \end{pmatrix}, R^7 = \begin{pmatrix} 6 & 5 & -1 \\ 0 & 5 & 1 \\ 6 & 3 & -2 \end{pmatrix}$$

$$P^8 = \begin{pmatrix} .2 & .5 & .3 \\ .1 & .6 & .3 \\ .05 & .4 & .55 \end{pmatrix}, R^8 = \begin{pmatrix} 7 & 6 & 3 \\ 7 & 4 & 0 \\ 6 & 3 & -2 \end{pmatrix}$$

Table 2.3

The values of V^k_i can thus be computed as given in the following table.

	V^s_i		
s	i = 1	i = 2	i = 3
1	5.3	3	-1
2	4.7	3.1	.4
3	4.7	3	-1
4	5.3	3.1	-1
5	5.3	3	.4
6	4.7	3.1	-1
7	4.7	3	.4
8	5.3	3.1	.4

The computations of the stationary probabilities are achieved by using the equations.

$$\pi^s P^s = \pi^s$$

$$\pi_1 + \pi_2 + \ldots + \pi_m = 1$$

Consider s = 2. The associated equations are

$$.3\pi_1 + .1\pi_2 + .05\pi_3 = \pi_1$$

$$.6\pi_1 + .6\pi_2 + .4\pi_3 = \pi_2$$

$$.1\pi_1 + .3\pi_2 + .55\pi_3 = \pi_3$$

$\pi_1 + \pi_2 + \pi_3 = \pi_1$

(Notice that one of the first three equations is redundant.).The solution yields as follow for policy 2.

$$\pi^2_1 = \frac{6}{59},\ \pi^2_2 = \frac{31}{59},\ \pi^2_3 = \frac{22}{59}$$

In this case, the expected yearly revenue is

$$E^2 = \sum_{i=1}^{3} \pi^2_j v^2_1$$

$$= \frac{1}{59}(6 \times 4.7 + 31 \times 3.1 + 22 \times .4) = 2.256$$

The following table summarizes π^k and E^k for all the stationary policies. (Although this will not effect the computations in any way, note that each of policies 1,3, 4, and 6 has an absorbing state: state 3. This is the reason $\pi_1 = \pi_2 = 0$ and $\pi_3 = 1$ for all these policies,)

Table 2.4

Policy 2 yields the largest expected yearly revenue.The optimum long-range policy calls for applying fertilizer regardless of the state of the system.

	V^s_i			
s	π^2_1	π^2_2	π^2_3	E^s
1	0	0	1	-1
2	$^6/_{59}$	$^{31}/_{59}$	$^{22}/_{59}$	2.256
3	0	0	1	.4
4	0	0	1	-1
5	$^5/_{154}$	$^{69}/_{154}$	$^{80}/_{154}$	1.724
6	0	0	1	-1
7	$^5/_{137}$	$^{62}/_{137}$	$^{80}/_{137}$	1.734
8	$^{12}/_{135}$	$^{69}/_{135}$	$^{54}/_{135}$	2.216

3 CONCLUSION

In the finite-stage model of the gardener's problem case1, the optimal solution shows that for years 1 and 2, the gardener should apply fertilizer (k* = 2) regardless of the state of the system. In year 3, fertilizer should be applied only if the system is in state 2 or 3 (fair or poor soil condition). The total reward for the three years are f_1 (1) = 10.74 if the state of the system in year 1 is good , f_1 (2) = 7.92 , if it is fair, and f_1 (3) = 4.23, if it is poor.

In the exhaustive enumeration method of infinite-stage model of the gardener's problem Case II, policy 2 yields the larges expected yearly revenue (2.256).

The optimum long – range policy calls for applying fertilizer regardless of the state of the system. In the alternative, contingency and reliability tests were performed and it was found that: the x^2 values of .936 to 6.146 were interpreted from x^2 table of probability values at 0.100 level of significance. The degree of freedom necessary to intercept x^2 values were determined from the frequency table by the number of rows minus one times number of columns minus one (r-1) (c-1) i.e (3-1)(3-1) = 4.

Since the obtained x^2 values of 0.936 to 6.146 were less than the critical value of 7.77944, the null hypothesis are accepted. There is relationship between the state of the system this year and the system state of the system next year. The chi square was not based on a fictitious data , for the case of the gardener's problem when fertilizer is applied and not applied. In the test of how well the linear estimator, y = a + bx fits the raw data, Correlation Coefficient, r which ranges from 0.4 to 1.0 result in good to perfect linear fit for the raw data. Conclusively, there is no significant difference between the experimental and theoretically expected result, which led to the acceptance of null hypothesis.

References

Hiller F.S and Lieberman G.J. (1995) "Introduction to Operations Research" Sixth Edition, Mc Graw – Hill, Inc America

Sharma J.K (2008) " Operations Research: Theory and Applications", Third Edition, Rajiv Beri for Macmillan India Ltd.

Anyata B.U (2009) "System Analysis and Numerical Methods in Civil Engineering, University of Benin City, Nigeria.

Ebeling C.E. (1997) " An Introduction to Reliability and Maintainability Engineering" Mc Graw – Hill, Boston.

Stamatelatos, M. et al., (2002) "Practitioners", NASA Office of Safety and Mission Assurance.

Vesely, W. et al., (2002) "Fault Tree Handbook with Aerospace Applications", NASA Office of Safety and Mission Assurance.

Chan T. et al (2010) "Participatory Development of a Bayasian Network Model for Catchments-based Water Resources Management," Water Resources Research, Vol. 46 N0 7544, pp.12

Danial P.L (2009) "Sustainability Implications for Water Resources planning and Management", A United Nations Sustainable Development Journal, vol. 18, Issue 4, pp.263 – 274.

Slobodan P.S. (2000) "Water Management and Sustainability", Water International, vol. 25, Issue 1 pp. 76-88.

Daniel P.L (2000) "Sustainable water Resources Management" Water International vol. 25, Issue 1, pp. 3-10.

Eme L.C (2004) "Application of Markovian Decision Theory in Multi-Purpose / Multi-Objective Water Resources planning and Management", Post Graduate project (M.Eng.), Nnamdi Azikiwe University Awka.

Peter N. (1980) , "An Attractiveness Analysis of Environmental Quality and Recreation", Environmental policy Analysis: Operational Method and models, John Miley & Sons, pp.262 – 263.

Nwaogazie I.L (2006) "Probability and Statistics for Science and Engineering practice", University of Port-Harcourt press.

Yi-bin et al (2004) " Multi-Person Multi-Objective Fuzzy Decision – Marking Model for Reservoir Flood Control Operation", Water Resources Management. Vol. 18 No. 2 pp. 111-124.

Lyupchop (2008) “Application of Contemporary Simulation Models for Management of Water Resources System”, University “Sts Cyril and Methodius” Faculty of Civil Engineering – Skope, Republic of Macedonia, ohrid.

John W.N (2000) “Discrete – time Optimal Control for Water Resources Engineering and Management” Water International, vol. 25 Issue 1,pp. 89-95.

Taha H. A. (2002) ”Operation Research an Introduction” Sixth Edition, Prentice Hall of India Private Limited, New Delhi-110 001.

Dr. Luke C. Eme is a Senior Lecturer and former Head of Department of Civil Engineering, Anambra State University, Uli-Nigeria. He holds First Degree, Masters of Engineering and Doctor of Philosophy Degrees in Civil Engineering with Distinctions and Specializes in Water Resources Engineering and Environmental Engineering. He had several publications and a member of many professional bodies and organizations which include: IRDI, IWA London, AASRC etc. He had received many prizes/awards nationally and internationally such that the outcome his Ph.D research work in 2012 at Anambra State University accorded him an International Award from Glen Daigger, President of the International Water Association London in the form of Books/Journals which he donated to the Library of Anambra State University. He received the honor for emerging first at a poster presentation held to mark the Association’s 2012 World Water Congress and Exhibition, during the competition held from September 16 to 21, at Bussan, Korea, a total of 387 poster presenters participated. Out of this number, five were short listed for the finals where Dr. Eme’s presentation was chosen as the best in the World.