UNIVERSITY OF SOUTHAMPTON

REPAIR STRATEGIES IN AN UNCERTAIN ENVIRONMENT

by

Yeek-Hyun Kim

Submitted for the degree of Doctor of Philosophy

School of Management
Faculty of Law, Arts and Social Sciences

March 2004

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF LAW, ARTS AND SOCIAL SCIENCES
SCHOOL OF MANAGEMENT

Doctor of Philosophy

REPAIR STRATEGIES IN AN UNCERTAIN ENVIRONMENT

By Yeek-Hyun Kim

This research deals with repair strategies which maximise the time until a catastrophic event - There is a vital need for the equipment, and it is unable to respond. We examine the case where the need for the equipment varies over time according to a Markov chain. This means that the environment can be in different states, each with their own probability of the initiating event occurring. We describe the form of the optimal policy under this uncertain environment by Markov Decision Process.

We also look at conflict situations where the environment is controlled by an opponent. In this case the opponent's actions force the need for the equipment, and this situation is modelled as a stochastic game. For this research, we develop stochastic game models with global and local constraints on effort. In the model with global constraints on effort, we introduce the idea of a constraint on the average effort undertaken by the opponent over the total history of the game so far. We naively describe this as a sleep index in that the opponent needs to sleep for a certain percentage of the time. We also expand these results to the situation where the advantage of a rest or quiescent period is discounted the further in the past it is, but always has a positive effect. In the model with local constraints on effort we look at games where the benefit to the opponent of being ' able to sleep' only lasts for a finite period and is then lost completely.

As extension of the first research, we also consider training. Because training is very important to increase the operator's responding ability against an initiating event, this new model is more realistic. However, a problem with training is that it increases wear and tear of stand-by units. We develop discrete time Markov decision process formulations for this problem in order to investigate the form of the optimal action policy.

CONTENTS

Chapter 3. Repair Strategies in an Uncertain Environment : Markov Decision Process Approach 35

Chapter 4. Repair Strategies in an Uncertain Environments : Stochastic Game Approach 72

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I am grateful to my God who gave me the ability to complete this research

I would like to express my most great gratitude to my supervisor, Professor Lyn C. Thomas, for his excellent supervision, untiring interest, remarkable help and encouragement for my research.

Many thanks go to the staffs and colleagues in the School of Management for making my years at the university a very memorable experience.

Thanks are also due to my sponsor, Republic of Korea Army, who supported me for this research. My special thanks go to COL Kwang-Mook Choi for his valuable advice for my Ph.D study.

Finally, I am greatly indebted to my parents, my wife, Hye-Jeong, and my lovely two children, Guhn-Hee and June-Hee, for their patience, understanding and support for me during the research period.

CHAPTER 1

Introduction

Our modern systems are composed of several types of equipment. If part of the equipment in the system fails, severe damage may occur. In order to avoid this situation, we install cold stand-by equipment which will improve the system's reliability. The stand-by units can be defined as an equipment which is only brought into operation when there is a vital need for it. For example, in hospital we may have a catastrophic situation if the main electric supply from the electric power company is not available for medical equipment. In order to overcome this severe situation all hospitals have their own emergency power supply system. This is a typical example of a stand-by unit. Military equipment including weapons is another example. Normally military hardware is reserved to respond to a situation which has real security dangers like enemy attack or war. If such a military stand-by unit fails to respond when it is required, the damage could be catastrophic. We call the times when there is a vital need for the stand-by unit as initiating events.

Even though there are many standby units in the civilian sector, we will concentrate mainly on the military standby units and circumstance in this research. If the environment around a stand-by unit changes, we may need to have different replacement and maintenance strategies for the stand-by unit. The military security environment is very uncertain, very flexible, easily changeable, especially during war time while the civilian environment is more stable, more certain, and less changeable. There has been a great deal of research in the maintenance and replacement policies for equipment, but, most of it has assumed a fixed, stable environment. This is not appropriate in the rapidly changing circumstances of a military environment. Hence it is useful to undertake research on the repair strategies for standby units under a changing environment.

This research is concerned with developing repair strategies which maximise the time until a catastrophic event for the standby unit in an uncertain environment. For this, we develop discrete time Markov decision process formulations of the standby unit maintenance problem. This allows us to calculate numerically the best policy in particular cases and to investigate the general form of the optimal maintenance policies which maximise the expected survival time until a catastrophic event. Although other approaches are possible, the Markov approach gives rise to models which are solvable, and in most cases the Markov application is a realistic one. The reason for focusing on the expected survival time rather than on cost is because we assume that the cost is immeasurable or very large if the system failed. In some senses, we are considering the operational policy separated from the purchase policy which is where the main costs are. So in operation in a military context the cost of failure is enormous either because of the consequences or because more units have to be purchased to overcome the possibility of one failing. In some ways, this criterion is like a discounted availability criterion. It balances the unavailability due to repair with the unavailability due to failure.

However unlike availability which is a steady state long term average, this criteria puts most emphasis on non-availability in the immediate future.

Markov decision processes are versatile and powerful tool for analyzing probabilistic sequential decision processes with a infinite planning horizon. They are an outgrowth of the Markov model and dynamic programming. Dynamic programming , developed by Bellman[1957], is a computational approach for analyzing sequential decision processes with a finite planing horizon. Part of our research is also about a game model in which the enemy controls the environment situation. In such cases, the state of readiness of troops or emergency services does fluctuate according to actions by the enemy country or the enemy terrorists. In such cases the opponent's actions affect the need for the stand-by equipment and hence its maintenance and repair schedule. We model such situations as a stochastic game. Our last problem for this research is about repair strategies where the failure of the system can be due both to its deterioration and to insufficient training of the operators on it.

In this thesis, we are ultimately concerned with the problem of repair strategies for a standby unit in changing environment. In order to study and solve the maintenance problem for standby units in a changing environment, first of all, we review the general research background literature in chapter 2, and develop the Markov decision process model with 1 repair action and with 2 repair actions under multiple environments in chapter 3. For the Markov decision process with 2 repair action model, we use two different repair which have different repair characteristics.

In chapter 4, using the stochastic game method, we look at the problem when there is conflict. For this game, we develop a basic model, and then we extend this basic model to more complicated models in order to overcome its limitation. Finally, in chapter 5, we consider another state space factor for our model, namely the training of the operator. Conclusions are presented in chapter 6.

CHAPTER 2

Literature Review

2.1 Introduction

Over the last few decades, many papers have appeared on literature which deal with the problem of finding optimal repair/replacement policies for systems which are subject to failures. One major influence in this field is traced back to Bellman's introduction of dynamic programming in 1957. Dynamic programming is a computational approach for analyzing sequential decision processes and is very useful when the stochastic elements are introduced in replacement problems. Also, in 1960 Howard developed a policy iteration method to solve multistage decision problems whose dynamics was a Markov process. Markov decision process combines the Markov models and dynamic programming. After Derman developed the original model of a single deteriorating unit in 1963 using Markov decision process, there has been a great deal of research in maintenance and replacement field. This maintenance and replacement phenomenon is indicated in various surveys by McCall[1965], Pierskalla[1976], Sherif and Smith[1982], Monahan[1982], Thomas[1986], Valdes-Flores and Feldman[1989], Cho and Parlar[1991], Dekker[1996] and Wang[2002] etc. However most of research concentrates on cost criterion in order to find the maintenance and replacement policy which minimises the cost per unit time. Few papers have considered expected survival time. In one such paper, Thomas et al[1987] developed a discrete time Markov decision process model for a standby system maximising the expected time until a catastrophic event. The maintenance and repair models considering random environment have been studied by Çinlar[1984], Çinlar and Özekici[1987], Çinlar et al[1989], Shaked and Shanthikumar[1989], Lefèvre and Milhaud[1990], Özekici[1995][1996], Wartman and Klutke[1994], Klutke et al[1996], Yang et al[2000a][2000b] and Kissler et al[2002]. Also, there have been several papers looking at reliability models under random environment by Dhillon et al[1982]etc, Goel et al[1985], Cao et al[1988] etc. Yeh[1995] studied an optimal maintenance model for a standby system but focusing on availability and reliability as the criteria to optimise.

The need to model the maintenance process as a game because an opponent is able to set the environment conditions has not been discussed before. In fact the application of game theory in the maintenance problem is restricted to the work of Murthy and Yeung[1995] and Murthy and Asgharizadeh[1999] in the case of warranty contracts. There has been no maintenance / replacement model using stochastic games. Here we will use the ideas of stochastic games to model our situation. Stochastic games were first introduced by Shapley[1953] who described the discounted game case almost precisely,while Gillette[1957], Blackwell and Ferguson[1968], Mertens and Neyman[1981] etc have looked at average cost game. In this chapter, the relevant literature on the above topics including Markov decision process, maintenance and replacement models, maintenance and replacement models for standby units, maintenance and replacement models under changing environment and stochastic games is reviewed.

2.2 Markov Decision Process

2.2.1 History The appearance of the Markov decision process is well documented by Putterman[1994], Heyman and Sobel[1984]. Bellman[1957] and Howard [1960] popularized the study of sequential decision processes. However we can go back to the calculus of variations problems of the 17th century for an early approach to such problems. Cayley's paper[1875] proposed a problem which contains many of the key ingredients of a stochastic sequential decision problem. The modern study of stochastic sequential decision problems began with the Wald's work[1947] on sequential statistical problems during the Second World War. He presented the essence of the Markov decision process theory in his research. Pierre Masse[1946], director of French electric companies and minister in charge of french electrical planning, introduced many of the basic Markov decision process concepts in his extensive analysis of water resource management model. A description of Masse's research is given in Gessford and Karlin[1958]. The modern foundations of Markov decision process were developed between 1949 and 1953 by people at RAND corporation, USA, for example, Bellman and Blackwell[1949];Bellman and LaSalle[1949];Shapley[1953] in games, Arrow et al[1951]; Dvoretsky et al[1952] in stochastic inventory models, Issacs[1955] in pursuit problems, Arrow et al[1949]; Robbins[1952]; Kiefer[1953] in sequential statistical problems. Among them Bellman was the first major player in the Markov decision process area. He identified common factors to these problems such as the functional equations of dynamic programming, and the principle of optimaility. Karlin[1955] also recognized and began studying the rich mathematical foundations of Markov decision process. Howard[1960] contributed to the progress in this field with his book on 'Dynamic Programming and Markov Processes' where he also introduced the so-called policy iteration method to solve multi-stage decision problems in connection with Markov processes.

2.2.2 Description of Markov Decision Process Markov decision process models consists of five elements: decision epochs, states, actions, transition probabilities, and rewards. When a system is being controlled over a period of time, a policy or strategy is required to determine what action to take in the light of what is known about the system at the time of choice, in terms of its state i. Decisions are made at points of time referred to as decision epochs. The state of the system describes the situation at a decision epoch. It may include information about past actions and events ,but it is called the state of the system at the time of the decision. In Markov decision process, if the decisions are made at uniform time intervals then, if the state is i at some decision epoch and if action a is taken, the system goes to state j at the next decision epoch with probability $P_{ij}(a)$. This is called the transition probability. In other words, let $s(k)$ be a Markov chain at time $k = 1, 2, - - -$, having dynamics described by the conditional probabilities

$$P_{ij}(a) = P[s(k+1) = j \mid s(k) = i,\ a(k) = a]$$

where $a(k)$ is the action taken at the decision epoch $k, a \in A$, $i, j \in S$. Let A be the finite action space, and S be the finite state space. Let n be the length of the time horizon. More generally, the action space can be dependent on the current state, $a(k) \in A[s(k)]$. The number of $A(i)$, actions in state i need not be equal to the number of $A(j)$. We can know that the transition probability from the current state of the process to the state of the process at the next decision epoch depends on the state i, j, and action a. The qualifier 'Markov' is used because the transition probability and reward functions depends on the past only through the current state of the system and the action selected by the decision maker in that state. The real valued reward $r_i(a)$ is accrued at decision epoch $k \leq n$ if $s(k) = i$ and $a(k) = a$. If the reward depends on the current state and action and the state at the next decision epoch, then we consider $r_i(a)$ to be the expected reward to be accrued until the next decision epoch, given the current state is i and the current action is a:

$$r_i(a) = \sum_{j \in S} r_{ij}(a) P_{ij}(a)$$

For the general problem with rewards $r_i(a)$ and transition probabilities $P_{ij}(a)$, the principle of optimality implies that the optimal total expected reward over n periods starting in state i, written $v^n(i)$, $n = 1, 2, ---, i \in S$ will satisfy the optimality equations:

$$v^n(i) = \max_{a \in A(i)} \{r_i(a) + \sum_{j \in S} P_{ij}(a) v^{n-1}(j)\}, \ i \in S, n = 1, 2, ---$$

This says that the optimal policy over n periods must after the first action be the optimal policy over the remaining $n-1$ periods. A rule is a contingency plan that selects actions and is described as a mapping $\delta : S \to A$. A policy is a sequence of rules $\pi = \{\delta_1, ---, \delta_n\}$, where δ_k is the rule to apply at decision epoch k. A policy or strategy specifies the decision rule to be used at all decision epoch. It provides the decision maker with a prescription for action selection under any possible future system state or history.

We do not have any problem to maximise or minimise the total expected rewards over a finite number of periods. However the trouble with infinite time horizon problems is that the total expected reward is infinite. In order to overcome this difficulty, we can use the following ways:

(1) discount all future rewards

(2) take average reward per period

If a reward in the first period is $r_i(a)$, the same reward in n th period is discounted as $\beta^{n-1}r_i(a)$ where $0 \leq \beta \leq 1$ is the discount factor. We can explain this by saying that future rewards are not worth as much as present rewards due to inflation. The optimality equation for maximising the total expected discounted rewards over an infinite horizon is

$$v(i) = \max_{a \in A(i)} \{r_i(a) + \beta \sum_{j \in S} P_{ij}(a)v(j)\}, \ i \in S$$

Hence, we say that the discounted reward over an infinite horizon is the reward over the first period plus the discounted reward from the second period onwards. Markov decision process is concerned with finding policies to produce optimal performance of the system to maximise the expected discounted reward ; or to maximise the reward per unit time or the expected total reward. Other objective may be appropriate to minimise the total expected cost or the long-run average cost per unit time. In our research, we have probability of an initiating event $b_m, 0 \leq b_m \leq 1$, according to each environment situation, $m, 1 \leq m \leq M$ where environment situation 1 is most peaceful situation and M is most dangerous situation. When the stand-by unit is in a non-operative state which occurs either in the down state or if it is being repaired, the stand-by unit will fail with certainty to respond an initiating event. The probability b_m of the initial event occurring depends on the state, m, of the environment. This means that in this case the stand-by unit can survive to the next time period with probability of $(1 - b_m)$. Since the value of $(1 - b_m)$ is located in the range of $0 \leq (1 - b_m) < 1$, this value acts like a discounted factor, β for our models.

In the average rewards case, if we let $v^n(i, \pi)$ be the total reward over n periods starting in state i, $i \in S$ under policy $\pi, \pi \in \Pi$, we define the average reward under policy π as $g(\pi)$ where

$$g(\pi) = \lim_{n \to \infty} \inf \frac{v^n(i, \pi)}{n}$$

In most cases, the total expected reward over n periods starting in state i, $i \in S$ under policy π, $v^n(i, \pi)$, converges to the form

$$v^n(i, \pi) \approx ng(\pi) + \omega(i, \pi)$$

where $\omega(i.\pi)$ is a biased value. We need to note that $\omega(i, \pi) - \omega(j, \pi) \approx v^n(i, \pi) - v^n(j, \pi)$ for n large, so that $\omega(i, \pi) - \omega(j, \pi)$ measures the difference in total expected costs when starting in state i rather than in state j, given that policy π is followed. If we assume that the optimal reward converges to $v^n(i) \approx ng + \omega(i)$ then g is the optimal average reward. If we substitute this into the finite horizon optimality equation, we can get

$$\begin{aligned} v^n(i) &= ng + \omega(i) \\ &= \max_{a \in A(i)} \{r_i(a) + \sum_{j \in S} P_{ij}(a) v^{n-1}(j)\} \\ &= \max_{a \in A(i)} \{r_i(a) + \sum_{j \in S} P_{ij}(a)[(n-1)g + \omega(j)]\}, \ i \in S, n = 1, 2, - - - \end{aligned}$$

Assuming $\sum_{j \in S} P_{ij}(a) = 1$, we have the optimality equation

$$g + \omega(i) = \max_{a \in A(i)} \{r_i(a) + \sum_{j \in S} P_{ij}(a)\omega(j)\}, \ i \in S$$

2.2.3 Solution of Markov Decision Process In general we have an optimal policy and the criterion value generated by this policy for each state in S as a solution to the standard Markov decision process. Procedures for determining a solution have been based on dynamic programing. Let $v^n(i)$ be the value of the criterion generated by an optimal policy assuming that $s(1) = i$ and that the horizon has length n. Then, the array $v^n = \{v^n(i) : i \in S\}$ can be determined from the recursion

$$v^n(i) = \max\{r_i(a) + \beta \sum_j P_{ij}(a) v^{n-1}(j) : a \in A\}$$

where $n = 1, 2, - - -$ and $v^0(i) = \bar{r}_i$. It is clear that δ_k^* is an optimal rule at decision epoch k if and only if it causes the maximum in above equation to be attained for all $i \in S$. This implies that when $s(k) = i$, $a(k)$ should be selected to equal $\delta_k^*(i)$. The sequence of such rules, $\pi^* = \{\delta_1^*, \delta_2^*, - - -, \delta_K^*\}$,constitutes an optimal policy.

In order to find the optimal policy and value, there exist several methods including value iteration method, policy iteration method, hybrid(modified policy) iteration methods, linear programming, and approximating the problem, using the structure. Among them we will concentrate on the value iteration method and policy iteration method. Computational comparison of value iteration and policy iteration algorithm for discounted Markov decision processes was well researched by Thomas et al[1983][1986]. Policy iteration method was introduced by Howard[1960]. The policy iteration algorithm works as follows:

Step 1. (initialization). Choose a policy π

Step 2. (policy evaluation step). For the current rule π, compute the unique solution to the optimaility equation, $v^n(\pi)$

Step 3. (policy improvement). Try to find a policy π' satisfying $v^n(\pi') > v^n(\pi)$

(i) If no such policy exists, π is optimal then STOP

(ii) If a π' can be found , put $\pi = \pi'$ and go to step 2.

The policy iteration algorithm is empirically found to be a remarkably robust algorithm that converges very fast in specific problems. The number of iteration is practically independent of the number of states. But the policy iteration algorithm requires that in each iteration a system of linear equations of the same size as the state space is solved. In general, this will be computationally burdensome for a large state space and this algorithm is computationally unattractive for large-scale Markov decision problems. So we need an alternative algorithm which avoids solving systems of linear equations but uses instead the recursive solution approach from dynamic programming. This method is the value-iteration algorithm which computes recursively a sequence of value functions approximating the optimal value. The value functions provide lower and upper bounds on the optimal value. In general value iteration algorithm endowed with these lower and upper bounds is the best computational method for solving large-scale Markov decision problems. It turns out that in value iteration the number of iterations is typically problem dependent. Another important advantage of value iteration is that it is usually easy to write an own code for specific applications. The value iteration algorithm computes recursively for $n = 1, 2, - - -$ the value function $v^n(i)$ from

$$v^n(i) = \max\{r_i(a) + \beta \sum_j P_{ij}(a) v^{n-1}(j) : a \in A\}, i \in S$$

starting with arbitrarily chosen function $v^0(i)$, $i \in S$. The quantity $v^n(i)$ can be interpreted as the maximal total expected rewards with n periods left to the time horizon when the current state is i and a terminal cost of $v^0(i)$ is incurred when the system ends up at state j. This interpretation suggests that the stationary policy whose actions maximise the right side of the above equation for all i will be very close to the minimal average rewards(see Tijms[1994]). We do not want to solve the optimality equation exactly, but we require ε-optimal solution v^* where $||v^* - v|| < \varepsilon$ and ε-optimal policies π^* where $||v^{\pi*} - v|| < \varepsilon$ and $||w|| = \sup_i |w(i)|$. For the optimal stopping of $v^n(i)$, we use McQueen's bound. McQueen's bound for the above value function is

$$v^n(i) + L(v^{n+1} - v^n)/(1 - \beta) \leq v(i) \leq v^n(i) + U(v^{n+1} - v^n)/(1 - \beta)$$

where $L(z) = \min_{i \in S} z(i)$ and $U(z) = \max_{i \in S} z(i)$ (see Thomas et al[1983]).

The value iteration algorithm works as follows(see Tijms[1994]):

Step 1. Choose $v^0(i)$ with $0 \leq v^0(i) \leq \min_a r_i(a)$ for all $i \in S$. Let $n := 1$

Step 2. Compute the value function $v^n(i), i \in S$,from

$$v^n(i) = \max\{r_i(a) + \beta \sum_j P_{ij}(a) v^{n-1}(j)\}$$

and determine $\pi(n)$ as a stationary policy whose actions maximise the right hand side of this equation for all $i \in S$.

Step 3. Compute the bounds

$$L(v^{n+1} - v^n) = \min_{i \in S}(v^{n+1} - v^n) \text{ and } U(v^{n+1} - v^n) = \max_{i \in S}(v^{n+1} - v^n)$$

The algorithm is stopped with policy $\pi(n)$ when $0 \leq U(v^{n+1} - v^n) - L(v^{n+1} - v^n) \leq (1-\beta)\varepsilon$ where ε is a prespecified tolerance number. Otherwise go to step 4.

Step 4. $n := n + 1$ and go to step 1.

The value iteration algorithm has in general not the robustness of the policy-iteration algorithm. The number of iterations required by the value-iteration algorithm is typically problem dependent and will usually increase when the number of states becomes larger. Also, the tolerance number ε in the stopping criterion will affect the number of iterations required.

McQueen's bound can also be used to eliminate non-optimal actions as well as to stop the iteration. McQueen[1967] showed that if

$$v^n(i,a) + \beta U(v^n - v^{n-1})/(1-\beta) < v^n(i) + \beta U(v^n - v^{n-1})/(1-\beta)$$

then the action a in $A(i)$ cannot be part of the optimal policy and hence can be removed from all future calculations. But Hastings and van Nunen[1977] noticed that we could modify the tests to check at the n th iteration whether action $a \in A(i)$ can be the maximising action for the $n+1$ th iteration. If it cannot, there is no point in calculating $v^{n+1}(i, a)$, so the action is removed temporarily from the calculations but then has to be tested again to see whether it can be the maximising action at the $n+2$ th iteration.

This action elimination method is very efficient to find an optimal solution for the cases which have many actions. In our Markov decision process model, we don't have to use the action elimination method since we do not have many actions to be considered. We have two actions, which are do nothing and repair, for the first Markov decision process model, and three actions, which are do nothing, quick repair and slow repair, for the second Markov decision model. In order to choose optimal action for each stage, we must know the expected survival time under each action. However, since the expected survival time under the repair action is independent of the previous quality state, we don't need to calculate this value at quality state. This means that once we have the expected survival time under repair for items in the new state, we can use this value in all the other states. Therefore we don't have to calculate the values under repair at each quality state. This means that we need to do only one calculation or at most two calculations. Hence, we do not have to use the action elimination method for this research.

2.3 Maintenance and Replacement Model

2.3.1 Original Model of a Single Deteriorating Unit by Derman

In order to understand the maintenance and replacement models, it is worth to review the first the original model of a single deteriorating unit described by Derman[1963]. A unit is inspected every period and the state of the unit is ascertained. It is assumed that if nothing is done then the unit deteriorates according to a Markov chain on a finite set of states $\{0, 1, 2, ---, L\}$, where 0 denotes a new unit and state L means the unit has failed completely. At each period, once the state is known, a decision has to be made whether to replace the unit, perform a preventive maintenance overhaul, or do nothing. The difference between replacement and repair is often a difference in which state the unit returns to after the performance of that action. In most variants of this model, economic criteria are used, in which one tries to minimise the sum of the maintenance cost, the cost of repair and replacement due to failure, and the cost of preventive repair and replacement. In this model, Derman showed that if the probability of deterioration next period increases with the present state i, then a "control limit" rule is optimal, so that one should repair or replace when the observed state i is greater than some limit i^*.

2.3.2 Extended Maintenance and Replacement Models

There have been a large number of extensions of this above original model, many of which are referred to in the excellent surveys of many authors. There are many possible ways to classify the works in maintenance and replacement models.

McCall[1965] surveyed scheduling policies for a stochastically failing equipment. In this survey, he says that maintenance models may be divided into two distinct categories. The first is the class of preparedness models in which the equipment fails stochastically and for at least some of its parts, its actual state is not known with certainty. Alternative maintenance actions for such equipment include inspection and replacement. Preventive maintenance models constitute the second class of maintenance models. In these models the equipment is subject to stochastic failure and the state of the equipment is always known with certainty. If the equipment exhibit an increasing failure rate and furthermore a failure in operation is more costly than replacement before failure, then, it may be advantageous to replace the equipment before failure. The problem is to determine a suitable replacement schedule. The basic division is between maintenance policies for which the equipment's failure distribution is known and maintenance policies for which the failure distribution is not known with certainty.

Pierskalla and Voelker[1976] suggest seven categories for maintenance and replacement model classifications. They say that one could establish a multi-dimensional grid whose coordinates would be (i) states of system, such as deterioration level, age, number of spares, number of units in service , number of state variables, etc., (ii) actions available, such as repair, replacement, opportunistic replacement, replacement of spares, continuous monitoring, discrete inspections, destructive inspections, etc., (iii) the time horizon involved, such as finite or infinite and discrete or continuous, (iv) knowledge of the system, such as complete knowledge or partial knowledge involving such things as noisy observation of the states, unknown costs, unknown failure distributions, etc., (v) stochastic or deterministic models, (vi) objectives of the system, such as minimise long run expected average costs-per-unit time, minimise expected total discounted costs, minimise total costs, etc., and (vii) methods of solution, such as liner programming, dynamic programming, generalized Lagrange multipliers, etc. In each cell of the grid, one could conceivably place every paper written on maintenance and replacement. However the work in this thesis would not fit within such a classification.

In Sherif and Smith's survey[1981] they suggest three classifications, ie., (i) general classification (ii) classification of maintenance models by type (iii) classification by type of applicable optimization technique. The general classification divides the maintenance and replacement models by whether they are inspection, maintenance, reliability, optimization techniques, or decision theory. For the second classification they divide optimal maintenance and repair models into deterministic models and stochastic models. Stochastic models can be divided into stochastic models under risk and stochastic models under uncertainty. Then these stochastic models under risk and stochastic models under uncertainty are divided into four different categories by whether they are simple or complex system, and use preventive maintenance or preparedness maintenance. In the last classification the optimization techniques employed for obtaining maintenance policies include the following: linear programming, nonlinear programming, dynamic programming, Pontryagin maximum principle, mixed-integer programming, decision theory, search technique, heuristic approaches.

Osaki and Nakagawa[1976] made a bibliography of reliability and availability of stochastic systems. This bibliography is directed towards methods using stochastic processes. In the bibliography, a brief review of stochastic models on system-reliability lists selected references on system reliability models using stochastic processes such as Markov chains, Markov processes, and semi-Markov(Markov renewal) processes. These stochastic process models provide decision criteria which aid in establishing optimum maintenance policies.

Thomas[1986] suggests next three types of maintenance and replacement models for multi-unit systems: (i) economic (ii) structural (iii) probabilistic. In economic type, the cost structure of replacement and maintenance has interdependencies between units. The simplest such case is when the replacement or repair cost of several components is less than the sum of their individual replacement or repair cost. In the structural type, it may be useful to replace working units at the same time as failed ones. Structural dependencies mean that one has to replace or at least dismantle some working units in order to replace or repair failed one. In probabilistic type, the state of one unit can affect the state of the other units or their failure rate.

Valdez-Flores and Feldman[1989] review preventive maintenance models where an optimal policy for a single-unit system(or a system that can be modeled as a single entry) is being determined. Although a system may consist of several components, it is sometimes practical to consider the system as a single unit that behaves in such a way that individual components do not directly affect the reliability of the system. Another important reason to consider single unit system is because in practice there are many instances in which it is difficult to obtain reliability data for smaller components; whereas data for the stochastic behavior of the entire system is available or easier to obtain. Maintenance models for single unit systems are also useful for modeling the maintenance of individual components that are part of more complex systems. In this survey the authors divide maintenance and replacement models into four categories: inspection models, minimal repair models, shock models, and miscellaneous replacement models. Inspection models are concerned with preparedness maintenance policies in which the current state of a system is not known but is available through inspection. Minimal repair models say there is a cost advantage from repairing a failed unit rather than from replacing it. They usually combine a periodic replacement policy with a minimal repair activity upon a unit failure. Shock models are concerned with single-unit systems where the unit is subject to exterior shocks which occur randomly.

Cho and Parlar[1991] survey the papers related to optimal maintenance and replacement models for multi-unit systems. In this survey, the papers are divided into five topical categories: machine interference/repair models, group/ block/ cannibalization/ opportunistic models, inventory/maintenance models, other maintenance/replacement models, and inspection/maintenance models. In machine interference/repair models, if any of the machines in a system interacts with one another a maintenance decision on one machine must be made in conjunction with what happens to the other machines in the system. Group/ block/ cannibalization/ opportunistic maintenance and replacement models are examined for the systems with identical unit. These models are mainly based on the concept of economies of scale. Inventory/maintenance models link the maintenance problems with the stockage problems.

Dekker[1996] reviewed applications yielding advice to management concerning maintenance on existing systems. Also he investigated what was needed for successful application of maintenance optimization models and what constituted the main bottlenecks. In his review, a literature search has been made on applications of maintenance optimization models and on tools developed to assist in maintenance optimization. Furthermore, discussions had been held with various people concerning the value of maintenance optimization models.

The maintenance of a deteriorating system is often imperfect. This means that the system after maintenance will not be as good as new, but younger than it was. Pham and Wang[1996] discussed and summarized various treatment methods and optimal policies of single and multi-component systems, and the rapidly growing literature on imperfect maintenance in order to give an overview of the recent maintenance models and policies, methodologies and techniques, tools, and applications. Also Wang[2002] surveyed more updated maintenance policies of deteriorating systems. This survey summarized, classified, and compared various existing maintenance policies for both single-unit and multi-unit systems. He also addressed relationships among different maintenance policies.

Christer[1982] addressed the idea of a "delay time" for a fault with the context of building maintenance. The main idea of the delay time concept is that defects do not just appear as failures, but are present for a while before becoming sufficiently obvious to be noticed and declared as failures. The time lapse from when a defect could first be reasonably expected to be identified at an inspection to the consequential failure repair if no corrective action is taken has been termed the delay time h of the default. The main key in delay time modelling is how to estimate the delay time distribution parameters. Since 1982, there have been a considerable numbers of papers published on inspection and maintenance modelling utilizing and developing the application of the delay time modelling, for example, Christer et al[2001], Pillay et al[2001], Wang[1997], Christer and Wang[1995], Baker and Christer[1994], Baker and Wang[1992][1993], Christer and Redmond[1990], Christer and Waller[1984a][1984b], etc.

2.3.3 Maintenance and Replacement Models for Standby Unit System reliability can be improved by providing standby units. Even a single standby unit plays an important role in a case that failure of an operating unit is costly and/or dangerous. As we mention before, a typical example of standby unit is a case of standby electric generators in hospitals and other public facilities. It is extremely serious if a standby generator fails when the electric power supply fails. Similar examples can be found in army defence systems, in which all weapons are standing by. Several works have been published on the problem of maintenance and replacement of standby units. Almost all previous work on standby units use a cost criterion.

Barlow and Proschan[1965] summarized schedules of inspections which minimise the two expected costs until detection of failure and per unit time assuming renewal of a standby unit. Luss and Kandar[1974] allowed for non-zero inspection times. Wattanpanom and Shaw[1979] studied the problem when the inspection is hazardous so that it is possible for the inspection to cause the unit to fail. Nakagawa[1980] looked at the probability that when there is a need for the standby units, the standby system will work. Butler[1979] maximised the expected lifetime of a standby unit, but did not allow for repairs. His model allowed the standby unit to be in more than one "up" state, which are distinguishable only upon inspection. This has a link with the problem of optimal inspection and repair of a deteriorating process with imperfect information introduced by Ross[1983] and generalized by White[1979], Rosenfield[1976], Luss[1976], Sengupta[1980], Suzuki[1979] and Wang[1977]. In these papers a system can be in more than one state, but which one is known only imperfectly or only upon inspection. The idea of using a catastrophic event criterion to overcome the problem that failure of stand-by units will result in unqualtifiably large cost was suggested first by Gaver, Jacobs and Thomas[1987]. In this model, discrete time Markov decision process models are formulated and policies for periodic inspection and maintenance of such units are derived to maximise the expected time until the standby unit can't respond on a need for it. First, they introduced a basic discrete time model where the unit can only be "up" or "down". The time between when there is a need for a standby unit are assumed to be independent random variables having common geometric distribution. In this basic model, they derive the form of the optimal policy which maximises the expected time until the standby unit can't respond on a need for it and study its behavior as a function of the parameters of the model. This yields information concerning the times to inspection after a repair and after a successful inspection. From this basic model the paper also consider a generalization of which allows the unit to be in one of two "up" states, which are indistinguishable on inspection, but which have different failure rates. This modification incorporates the idea that the repair might deal with the superficial cause of the unit's failure, but miss the underlying problem which recur. In certain cases, the optimal inspection policy for this model has quite short inspection periods immediately after a repair which then lengthen as further inspections suggest that the unit is in better "up" state. Yeh[1995] studied an optimal maintenance model for standby system focusing availability and reliability as the criteria to optimise. In his paper, an inspection-repair-replacement policy is employed. He assumes that the state of the system can only be determined through an inspection which may incorrectly identify the system state. After each inspection, if the system is identified as in the down state, a repair action will be taken. It will be replaced some time later by a new and identical one. So he can developed an optimal policy so that the availability of the system is high enough at any time and the long-run expected cost per unit time is minimised.

2.3.4 Maintenance and Replace Models under changing Environment The maintenance and replace models considering random environment have been considered by several people. It is natural to assume that the level of environmental exposure changes randomly over time, and hence that the deterioration rate can be modeled as a stochastic processes. The presence of a random environment adds considerable complexity to these models. Çinlar[1984], Çinlar and Özekici[1987], Özekici[1995][1996] studied several models of Markovian and semi-Markovian deterioration in the context of determining a single device lifetime. The essential idea used in their analysis is the correspondence between 'real time(in actual use under field conditions)' and an 'intrinsic age process' which measures time in units of deterioration. In Çinlar et al[1989] and, independently, Lefèvre and Milhaud[1990] they consider the lifelengths $T_1, ---, T_k$ of k components subjected to a random varying environment. They are dependent on each other because of their common dependence of the environment. The parameters of the model are the distribution of the random process which describes the environment and a set of rate functions which determine the probability law of $T_1, ---, T_k$ as a function of the distribution of the environment. They find conditions on the parameters of the model which imply that $T_1, ---, T_k$ are associated. Shaked and Shanthikumar[1989] provide more precise models from Çinlar and Özekici[1987] and Çinlar et al[1989]. Because the hazard rate of a component may be given directly as a function of the environmental state, the present time, and the previous failure time not by differential equation, they describe such a replacement model which gives rise to dependence of the lifetimes of the different components among themselves through their dependence on the common environment. They also describe the monotonicity of various probabilistic quantities of interest as functionals of the distribution law of the environmental process. But in all these cases the unit(system) is always in use so the changes in the environment age the equipment at different rates, but do not affect when it is needed.

Many protective systems, such as circuit breakers, alarms, and protective relays, as well as spare or standby systems have non-self-announcing failures where the rate of deterioration is governed by a random environment. These systems have been studied by Wartman and Klutke[1994], Klutke et al[1996], Yang et al[2000a][2000b] and Kissler et al[2002] while in our research we consider the fixed rate of deterioration with different probability of need for the standby equipment along the environment. There are some papers on reliability under changing environment.

Dhillon et al[1982] studied the behaviour of a single unit system under two different weather condition - normal and stormy. In this paper, they presented two mathematical Markov models to predict human reliability of continuous operating tasks. Dhillon et al[1985][1986] showed multi mathematical Markov models to predict electric power systems reliability in changing environment(normal and stormy weather).

Goel et al[1985-1] pointed out difficulties in the models of Dhillon et al[1982] and proposed a revised version of their models with two modes, i.e. normal and total failure. In Goel et al [1985-2] they consider three modes, i.e. normal, partial failure and total failure under two weather condition(normal, stormy). Then the cost analysis of a system having these three modes are considered. The paper by Goel et al[1985-3] deals with the cost analysis of a two unit cold standby system different weather conditions(normal and abnormal).

In other papers which have analyzed repairable systems in different weather conditions, the rates of change of weather conditions are assumed to be constant. The paper by Cao[1988-1] considers a repairable system in a changing environment subject to a general alternating renewal process. He obtains the system availability, failure frequency and reliability function. The paper by Cao[1988-2] discusses a man-machine system operating under changing environment where the change of the environment is subject to a Markov process with two state. Using Markov renewal process Cao obtains the system availability, future frequency and reliability function.

Guo and Cao[1992] consider a one-unit repairable system operating under changing environment and assume that the system has m types of failure under one environment and n types under another. The rates of change of environment and failure are constant while the repair rates are general. By using Markov renewal theory, they obtain the system availability, failure frequency and reliability function.

Lien[1992] considers a repairable system which operates in a random environment. The changes of environmental levels are described by a semi-Markov process with finite states. He assume that both this system's life time and repair times have exponential distribution but that their parameters change with environmental levels. The distribution of the first up and down time of the system and the system availability are obtained. By an alternating renewal process, he also get the bounds of the system first failure mean time.

Song and Deng[1993-1] consider a parallel redundant repairable system consisting of three identical units and one repair facility, which operates in a changing environment subject to a Markov process with two states. Using Markov renewal process, they get the system availability, failure frequency and reliability function. Song and Deng[1993-2] consider a system which operates in randomly changing environment. The changes of environmental levels are described by a Markov process with finite states. The system consists of m identical units and one repair facility. When the system operates in environment level i and has j failed units, the working units of the system has a failure rate λ_{ij}. The repair time of a failed unit has an arbitrary distribution. Song and Deng obtain the system reliability functions, availabilities and failure frequencies.

2.4 Stochastic Game

2.4.1 History

Game theory was created by von Neumann in his two papers[1928, 1937]. However, it came to life in the book, 'Theory of Games and Economic Behaviour' written by von Neumann and Morgenstern[1944]. In their book, von Neumann and Morgenstern investigated two distinct possible approaches to the game theory. The first of these is the noncooperative approach. This requires specifying in close detail what the players can and cannot do during the game, and then searching for a strategy for each player that is optimal. What is best for one player depends on what the other players are planning to do, and this in turn depends on what they think the first player will do. von Neumann and Morgenstern solved this problem for the particular case of two-player games in which the players' interests are diametically opposed. Such a game is called zero-sum game because any gain by one player is always exactly balanced by a corresponding loss by the other. In the second part of the book, von Neumann and Morgenstern developed cooperative approach, in which they sought to describe optimal behavior in games with many players. Due to the difficulty of the problem, they did not attempt to specify optimal strategies for individual players. Instead they look to classify the patterns of coalition formation. In a sequence of remarkable papers in the fifties, the mathematician John Nash found two important discoveries breaking the restriction that von Neuman and Morgenstern had erected for themselves. In noncooperative game, they felt that the idea of an equilibrium in strategies was not an proper notion and hence their restriction to zero-sum games. However Nash's general formulation of the equilibrium idea made it clear that no such restriction is necessary. Nash also contributed to von Neumann and Morgenstern's cooperative approach. He disagreed with the view that game theorist must regard two-person bargaining problems as indeterminate, and proceeded to offer arguments for determining them[Binmore, 1992].

A stochastic game is one in which the outcome can be a real payoff and a requirement to play the same or another game again. The historic background of stochastic game is well explained in the book of 'Games, Theory and Applications' by Thomas[1986]. The concept of a stochastic game was introduced by Shapley[1953] in 1953 even though the idea of a dynamic game was used as a proof technique by von Neuman and Morgenstern[1947]. He introduced discounted stochastic games and gave a proof of the existence of a value for such game. Since Shapley's paper there have been new solution algorithms suggested, like the variation on value iteration given by van der Wal[1977], or ones called policy iteration which try to improve the strategies used at each iteration in Rao et al[1973]. For many years it was an open question whether or not the average reward per stage stochastic games had a value or not. Gillette[1957] proved the existence of the value in two cases: first when all games have perfect information and also in the so called cyclic case. Blackwell and Ferguson[1968] found in a particular example("Big match") two strategies that would prove to be basic for further generalizations. Bewley and Kohlberg[1976] made a large step on the road to solving the problem of existence of a value. Metens and Neyman[1981] proved the existence of a value for average reward stochastic games.

<u>2.4.2 Description of Stochastic Game</u> A stochastic game is one in which the outcome can be a real payoff and a requirement to play the same or another game again. Such a game may be regarded as a sequence of two-person zero-sum games played consequently, with the outcome of each game determining, stochastically, the matrix of the next. A two-person zero-sum stochastic game , Γ, is a set of N subgames, Γ_1, Γ_2, $---$, Γ_N. The normal form of subgame Γ_k is an $n_k \times m_k$

payoff matrix whose entries are of the form:

$$e_{ij}^{k} = a_{ij}^{k} \text{ and } \sum_{l=1}^{N} p_{ij}^{kl} \Gamma_l$$

where e_{ij} is the payoff if player I plays I_i strategy and player II plays II_j strategy. It consists of a numerical reward a_{ij}^{k} and$\sum_{l=1}^{N} p_{ij}^{kl}\Gamma_l$ where we assume $|a_{ij}^{k}| \leq M$, and a probability of playing Γ_l for $l = 1, 2, ---, N$, where $p_{ij}^{kl} \geq 0$ and $\sum_{l=1}^{N} p_{ij}^{kl} \leq 1$. Each time we play one of the subgames it constitutes a stage of the game Γ. If $\sum_{l=1}^{N} p_{ij}^{kl} < 1$, then there is a positive probability $1 - \sum_{l=1}^{N} p_{ij}^{kl}$ that the game ends at this stage. The difficulty with such general stochastic game is that (1) they could lead to unbounded payoffs, and (2) we cannot compare two strategies that both have such payoffs. To overcome this difficulty, we can consider two ways of thinking about the rewards from the game: (1) Discounting the payoffs, and (2) Taking the average reward per stage. In discounting the payoffs, the payoff from game played at r th stage is discounted by β^{r-1}, where $\beta \leq 1$. This is the same idea as inflation. Therefore a reward of 1 next stage is equal to a reward of β now. And a reward of 1 next two stages is equal to a reward of β^2 now. In taking the average reward per stage, if E_s is the reward from the first s stage of the game for a strategy, we look at $\lim_{s\to\infty} \frac{E_s}{s}$ as the average reward for that strategy. There is no good solution algorithm for stochastic game so far. However, Shapley[1953] and Avi-Itzhak[1969] found that discounted stochastic games always have solutions while Blackwell and Ferguson[1968] found that by taking average cost we cannot always solve stochastic games.

In discounted stochastic games a typical strategy for a player is very complicated. For the player I this strategy consists of a collection of strategies,

$$\{x^k(t, h_1, h_2, ---, h_{t-1}),\ k = 1, 2, ---, N,\ t = 1, 2, ---\}$$

where x^k is a (possibly mixed) strategy in the subgame Γ_k and depends on stage t, he is playing, and on the history $h_1, h_2, ---, h_{t-1}$ of what happened in the first $t-1$ stages. And h_i tell us information including which game was played at the ith stage and which strategy each player played in it. If $x^k(t, h_1, h_2, ---, h_{t-1}) = x^k$, for all t, k, and all possible histories $h_1, h_2, ---, h_{t-1}$, this is called a stationary Markov strategy. Even though each subgame Γ_k of the stochastic game Γ has a finite number of pure strategies, the overall game Γ has an infinite number of such strategies since one chooses an action for each subgame at each of an infinite number of stage. Hence the von Neumann minimax theorem(for a finite number of pure strategies) does not guarantee that the stochastic game Γ has a solution.

To prove that the stochastic game Γ has a solution, we need to use some trick. If the game has a value, it will depend on which subgame we started playing at the first stage. The value $v^*$of Γ(if it exists) is vector $\mathrm{v}^* = (v_1^*, v_2^*, ---, v_N^*)$, where $v_k^* = \mathrm{val}(\Gamma\ |\text{starting in } \Gamma_k \text{ at stage one})$. If we substitute the value of Γ starting in Γ_l instead of Γ_l, new payoff entries for this new game is $a_{ij}^k + \sum_{l=1}^N p_{ij}^{kl} v_l^*$. If we start Γ in Γ_k, then if we go to Γ_l at the second stage, the value of the game from there from there on is not v_l^* but βv_l^*. We are not playing Γ starting at Γ_l at the first stage but playing Γ with Γ_l at the second stage and so all the payoffs will be one stage later than under Γ starting in Γ_l. Hence, we can change payoff entries to discounted one:

$$e^k_{ij} = a^k_{ij} + \beta \sum_{l=1}^{N} p^{kl}_{ij} v_l$$

where $v_k = \text{val}(\Gamma_k(v))$, $k = 1, 2, - - -, N$ and $\Gamma_k(v)$ is an $n_k \times m_k$ game with payoff matrix. In general, we write $\Gamma_k(w)$ for an $n_k \times m_k$ game with payoff matrix[Thomas, 1986]:

$$e^k_{ij} = a^k_{ij} + \beta \sum_{l=1}^{N} p^{kl}_{ij} w_l$$

In our stochastic game model, we use a discretization technique for a continuous state space. For the decayed capacity limit model of this stochastic game models, we have to do approximation method in order to get sleep. Using this method we can make the continuous value of the sleep index into a set of discrete one.

CHAPTER 3

Repair Strategies in an Uncertain Environment: Markov Decision Process Approach

3.1 Introduction

This research is concerned with developing repair strategies which maximise the time until a catastrophic event for the standby unit in an uncertain environment. These environments can be thought of as the states DEFCON 1 to 5 or black, amber, red used by the US and UK military. For this, we develop discrete time Markov decision process formulations of the standby unit repair problem in order to investigate the form of the optimal repair policies which maximise the expected survival time until a catastrophic event. The reason for focusing on the expected survival time rather than on cost is because we assume that the cost is immeasurable if the system failed. We develop a Markov decision process model with one repair action under changing environment in section 3.2 and two repair actions under changing environment in section 3.4. For the Markov decision process with two repair action model, we use two different repair actions which have different repair characteristics. For each model we show a numerical example. Conclusions are presented in section 3.6.

3.2 Model for 1 Repair Action under Multiple Environment

3.2.1 Introduction

In this model, there are several environment situations which are graded from very dangerous to completely peaceful. Each environment situation has its own probability of the initiating event occurrence. There is only one repair maintenance action for this model which means that the operator can choose his best maintenance action between doing nothing and repairing.

3.2.2 Terminology

Possible Standby Unit Quality State, i Regular inspection of the standby unit gives information on the operation quality state of the units. We assume the standby unit has N different unit quality states, ie. $1, 2, - - -, N$ where state 1 means that the standby unit is like new. The state $N-1$ means that it is in a poor but still operable state, while in state N, it is in a "down" condition which means that it will not work. In the military context, there is often such a classification used into, new/excellent, operating, or failed, for example.

The Quality State Transition Probability Matrix(QSTPM), P_{ij} When the standby unit is in quality state i at the current stage, there is a probability, P_{ij} that it will be in state j at the next period where $i, j = 1, 2, \cdots, N$ and

$$\sum_{j=1}^{N} P_{ij} = 1, \text{ where } i = 1, 2, \cdots, N-1, N$$

We assume that the QSTPM satisfies a first order stochastic ordering condition so that $\sum_{j<k} P_{ij} > \sum_{j<k} P_{(i+1)j}$. We assume $P_{NN} = 1$ so once the standby unit

reaches the "down" state N, it remains "down" until either it is repaired, or a catastrophic event occurs. The value P_{ij} can be observed in practice from data collected in previous use of the equipment.

Possible Environmental Situation, m We assume that there are M different environmental states, $1, 2, \cdots, M-1, M$. Environmental state 1 reflects the most peaceful environment in which there is the smallest probability, b_1 of an initiating event occurring. On the other hand, the environmental state M is the most dangerous state with the highest probability, b_M of an initiating event occurring. We assume b_m is non-decreasing in the index of the environmental state m and $0 < b_m \leq 1$.

Environment Situation Transition Probability Matrix(ESTPM), $S_{mm'}$ The dynamics of the environmental situation is also described by a Markov chain with Environment Situation Transition Probability Matrix(ESTPM), $S_{mm'}$. If the environmental situation is m, $1 \leq m \leq M$ in the current stage, this changes to another environmental situation $m', 1 \leq m' \leq M$ with probability $S_{mm'}$ at the next stage, where

$$\sum_{m'=1}^{M} S_{mm'} = 1, \text{ with } m \text{ and } m' = 1, 2, \cdots, M-1, M$$

We assume the ESTPM also satisfies a first order stochastic ordering property so $\sum_{m'=1}^{k} S_{mm'} > \sum_{m'=1}^{k} S_{(m+1)m'}$ for any $m = 1, 2, \cdots, M-1$. A $S_{mm'}$ can be collected using historical experience on the changes in alertness levels.

The Possible Repair Actions There are two possible actions at each period, (1) do nothing and (2) repair. The "do nothing" action means that the operator does not repair the standby units. It is assumed the "repair" action takes 1 unit time period. Thus the time of the repair, including collecting and delivering, defines the unit of time generally used in those models. The repair action is not perfect in that there is a probability R_r the unit will be in quality state r after the repair where $\sum_{r=1}^{N} R_r = 1$. If an initiating event occurs during repair period, the standby unit cannot respond to it, and so a catastrophic event occurs automatically.

3.2.3 Model

Model State Space The state space of this model S has two factors which are unit quality state and environment state, so

$$S = \{(i,\ m) \in S,\ i = 1, 2, \cdots, N-1,\ N \text{ and } m = 1, 2, \cdots, M-1, M\}$$

where i and m mean the unit quality state and the environment situation respectively.

Maximum Expected Period, $V(i, m)$ When the unit is in quality state i and the environment situation is in state m, $V(i, m)$ is the maximum expected number of periods until a catastrophic event occurs.

Model Description In order to calculate the maximum expected period before a catastrophic event, we need to consider all cases making this model. For doing that, we need to think of each model from environment situation 1(most peaceful situation) to environment situation M(most dangerous situation). Then, for the each environment situation, we can consider each case of the "do nothing" maintenance action and "repair" maintenance action. When an initiating event occurs, the "do nothing" action can be divided by two sub-cases, (1) a case in which the standby unit quality state is in working condition($i \neq N$), and (2) a case in which the standby unit quality state is in N(failure state).

Case 1. Do Nothing
(1) When the standby unit is working, $i \neq N$
(2) When the standby unit is down, $i = N$
Case 2. Repair

Hence, we need to develop in total three models for each environment situation, (1) a model for the maximum expected period for the "do nothing" case when the unit quality state is in working condition ($i \neq N$), $W_1(i, m)$, (2) a model for the maximum expected period for the "do nothing" case when the unit quality state is in failure condition($i = N$), $W_1(N, m)$, and (3) a model for the maximum expected period for the "repair" case in which we do not need to consider the standby unit quality state, $W_2(m)$. Then, we can write an optimality equation to decide which expected period between $W_1$and W_2 by each action is longer than the other's. We can then decide the best maintenance action according to time, unit quality state and environment situation.

<u>Expected Survival Period by Do Nothing under Environment Situation m, $W_1(i, m)$</u>

In this case, the probability of an initiating event is b_m. We can consider the following two cases.

(1) When the standby unit is in working condition($i \neq N$ **)**

In this case, since the standby unit quality state is in various working conditions($i \neq N$), the unit does not have any failure to respond to an initiating event. The expected period of this case is not effected by occurrence of an initiating event during the next 1 time period between t and $t+1$. Hence, the state i moves on to the state j in the next stage with probability P_{ij} where $\sum_{j=1}^{N} P_{ij} = 1$. Also the environment situation goes from situation m to situation m' with a probability of $S_{mm'}$ where $\sum_{m'=1}^{M} S_{mm'} = 1$. Hence, the expected survival period by do nothing action when $i \neq N$ and m, $W_1(i,m)$ is

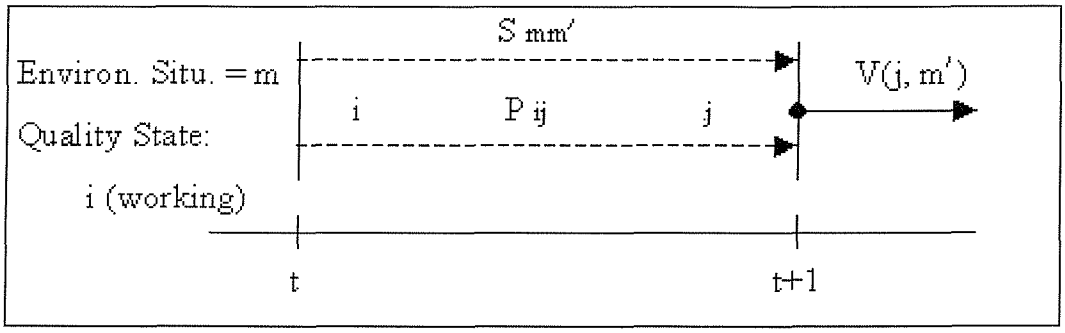

Figure 3-1. When the standby unit is in working condition

$$W_1(i,m) = 1 + \sum_{j=1}^{N} P_{ij} \sum_{m'=1}^{M} S_{mm'} V_{n-1}(j, m')$$

(2) When the standby unit is in failure condition($i = N$**)**

In this case the standby unit is in a vulnerable quality state because the unit is in the "down" state. Thus, if an initiating event occurs during the next 1 unit time period between t and $t+1$, a catastrophic event arrives automatically. When there is no occurrence of initiating event during the next 1 unit time period, the quality state of the unit still remains in the state of N, another vulnerable state at the next stage, $t+1$.

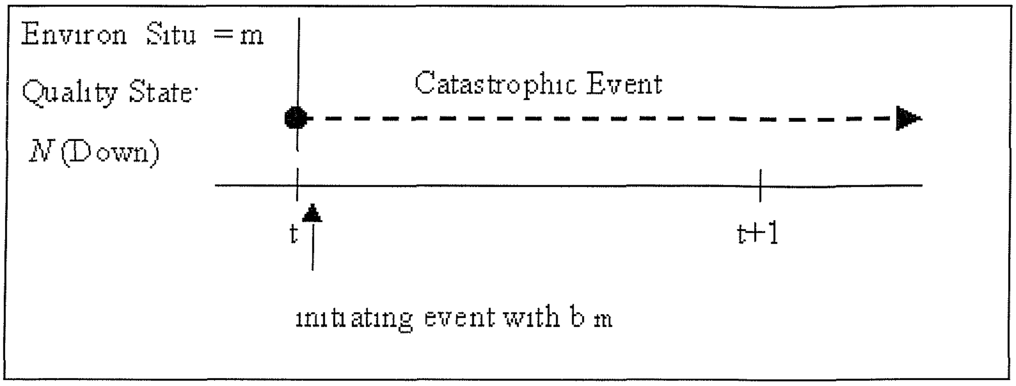

Figure 3-2. When the standby unit is in failure condition with an initiating event

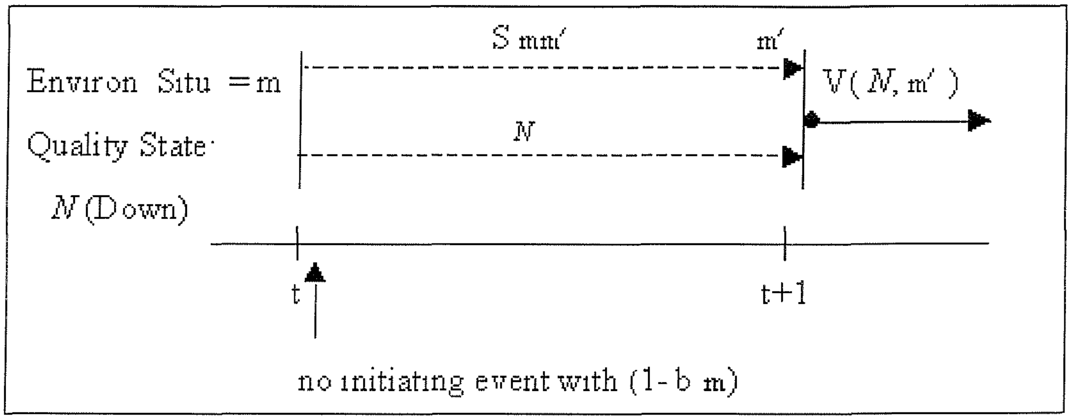

Figure 3-3. When the standby unit is in failure condition without an initiating event

Hence, the expected survival period by do nothing action when $i = N$ and m, $W_1(N,m)$ is

$$W_1(N,m) = (1-b_m)[1+\sum_{m'=1}^{M} S_{mm'}V_{n-1}(N,m')]$$

<u>Expected Survival Periods by Repair under Environment Situation m, $W_2(m)$</u> If the standby unit is in "repair" period which takes 1 time unit period, this is also another vulnerable period. If there is the occurrence of an initiating event, the

system has a catastrophic event automatically. Current quality state of the unit does not affect the occurrence of a catastrophic event since the unit is in repairing condition regardless of current quality state. We can consider this repair situation in two possible sub-cases, (1) when an initiating event occurs during the repair period, (2) when no initiating event occurs during the repair period.

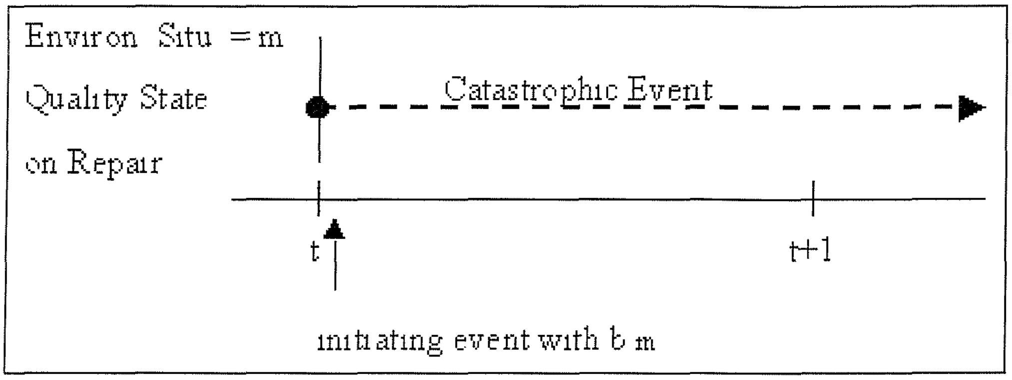

Figure 3-4. When the standby unit is in repair with an initiating event

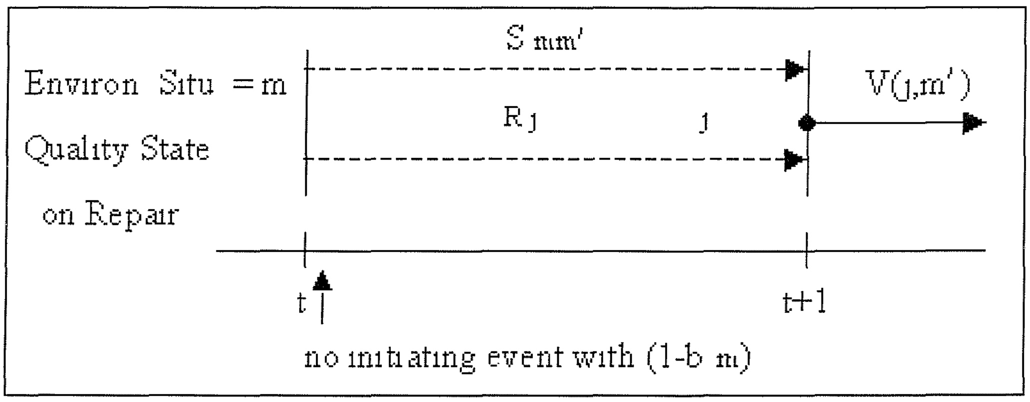

Figure 3-5. When the standby unit is in repair without an initiating event

Hence, the expected survival period by repairing when the environment situation is m, $W_2(m)$ is

$$W_2(m) = (1 - b_m)[1 + \sum_{m'=1}^{M} S_{mm'} \sum_{r=1}^{N} R_r V_{n-1}(r, m')]$$

Optimality Equation Because we are looking for the best repair policy, the optimality equation selects the better of the two actions.

$$V(i,m) = \max\{W_1(i,m), W_2(m)\} \tag{3.1}$$

,where $W_1(i,m)$ is the expected period until a catastrophic event if nothing is done now and $W_2(m)$ is the expected period until a catastrophic event if a repair is performed now. Hence the $W_1(i,m), W_2(m)$ can be expressed by

$$W_1(i,m) = (1 - \delta_{iN} b_m)[1 + \sum_{j=1}^{N} P_{ij} \sum_{m'=1}^{M} S_{mm'} V(j,m')] \tag{3.2}$$

where

$$\delta_{iN} = 1 \text{ if } i = N$$

$$\delta_{iN} = 0 \text{ otherwise}$$

$$W_2(m) = (1 - b_m)[1 + \sum_{m'=1}^{M} S_{mm'} \sum_{r=1}^{N} R_r V(r,m')] \tag{3.3}$$

$(3.1), (3.2), (3.3)$ can be solved by value iteration where the nth iterate satisfies

$$V_n(i,m) = \max\{W_n^1(i,m), W_n^2(m)\} \tag{3.4}$$

where

$$W_n^1(i,m) = (1 - \delta_{iN} b_m)[1 + \sum_{j=1}^{N} P_{ij} \sum_{m'=1}^{M} S_{mm'} V_{n-1}(j,m')] \tag{3.5}$$

$$W_n^2(m) = (1 - b_m)[1 + \sum_{m'=1}^{M} S_{mm'} \sum_{r=1}^{N} R_r V_{n-1}(r, m')] \tag{3.6}$$

If we define the terminal value, $V_0(i, m) = 0$, the $V_n(i, m)$ is a bounded increasing sequence of function and so converges to the limit $V_n(i, m)$. Standard results from Markov decision processes[Putteran,1994] show that the limit function satisfies the optimality equation $(3.1) - (3.3)$.

Lemma 3.2.1 $V(i, m)$ is a non-increasing function of i, $V(i, m) \geq V(i+1, m)$ where i is the quality state, $1 \leq i \leq N$, and m is the arbitrary environment situation state, $1 \leq m \leq M$.

(Proof)

The proof uses an induction hypothesis on n in $V_n(i, m)$. Trivially the property holds for $V_0(i, m)$. We can assume $V_{n-1}(i, m)$ is non-increasing function in i. This means $V_{n-1}(i, m) \geq V_{n-1}(i + 1, m)$. Also, from the stochastic ordering condition of QSTPM, we know $\sum_{j \leq k} P_{ij} \geq \sum_{j \leq k} P_{(i+1)j}$. And so, for (3.5)

$$\sum P_{ij} \sum_{m'=1}^{M} S_{mm'} V_{n-1}(j, m') > \sum P_{(i+1)j} \sum_{m'=1}^{M} S_{mm'} V_{n-1}(j, m')$$

Hence, we can conclude that $W_n^1(i, m) \geq W_n^1(i + 1, m)$. Since $W_n^2(i, m) = W_n^2(i+1, m)$, it follows $V_n(i, m) \geq V_n(i+1, m)$. Hence the result holds for $V_n(i, m)$ and by convergence the results hold in the limit for $V(i, m)$.

Lemma 3.2.2 $V(i, l) \geq V(i, l')$ for all i, where $l \leq l'$

(Proof)

The proof is similar to that of lemma 3.2.1 proving the result by induction in n on $V_n(i,l)$. The induction step follows from the stochastic order of ESTPM, S_{lm}. Hence, we can prove $V(i,l) \geq V(i,l')$ for all $i, l \leq l'$

<u>Theorem 3.2.1</u> In quality state N(down), one should always repair.

(Proof)

If $i^*(m)$ is given by $i^*(m) = \min[i : W_1(i,m) \leq W_2(m)]$ defined by (3.5) and (3.6), this theorem means $i^*(1), i^*(2), \cdots, i^*(M) \leq N$. We can let l be an arbitrary situation from the states $1, 2, \cdots, M$. In state N, the optimality equation is

$$
\begin{aligned}
V(N,l) &= \max\{W_1(N,l), W_2(N,l)\}, \text{where} \\
W_1(N,l) &= (1-b_l)[1 + \sum_{m'=1}^{M} S_{lm'} V(N,m')] \\
W_2(l) &= (1-b_l)[1 + \sum_{m'=1}^{M} S_{lm'} \sum_{r=1}^{N} R_r V(r,m')]
\end{aligned}
$$

If $W_2(l) \geq W_1(N,l)$, then $i^*(l) \leq N$.

$$
\begin{aligned}
& W_2(l) - W_1(N,l) \\
= \ & (1-b_l)[1 + \sum_{m'=1}^{M} S_{lm'} \sum_{r=1}^{N} R_r V(r,m')] - (1-b_l)[1 + \sum_{m'=1}^{M} S_{lm'} V(N,m')] \\
= \ & (1-b_l) \sum_{m'=1}^{M} S_{lm'} [\sum_{r=1}^{N} R_r V(r,m') - V(N,m')] \\
= \ & (1-b_l) \sum_{m'=1}^{M} S_{lm'} [\sum_{r=1}^{N} R_r V(r,m') - \sum_{r=1}^{N} R_r V(N,m')] \\
= \ & (1-b_l) \sum_{m'=1}^{M} S_{lm'} \{\sum_{r=1}^{N-1} R_r [V(r,m') - V(N,m')]\}
\end{aligned}
$$

Since $(1-b_l) \geq 0$, and $V(j,m)-V(N,m) \geq 0, r < N$, then $W_2(l)-W_1(N,l) \geq 0$. This means that , at quality state N, we should always repair : $i^*(l) \leq N$.

Theorem 3.2.2 The optimal policy is

1) repair in (i,m) if $i \geq i^*(m)$

2) do nothing if $i < i^*(m)$

(Proof)

The optimality equation is given by

$$\begin{aligned} & V(i,l) \\ = \quad & \max\{1+\sum_{j=1}^{N} P_{ij} \sum_{m'=1}^{M} S_{lm'} V(j,m'), (1-b_l)[1+\sum_{m'=1}^{M} S_{lm'} \sum_{r=1}^{N} R_r V(r,m')]\} \end{aligned}$$

From the above optimality equation, we can say that, it is optimal to repair the equipment in i if

$$(1-b_l)[1+\sum_{m'=1}^{M} S_{lm'} \sum_{r=1}^{N} R_r V(r,m')] > 1+\sum_{j=1}^{N} P_{ij} \sum_{m'=1}^{M} S_{lm'} V(j,m')$$

From the previous lemmas, we know that $V(i,l)$ is a non-increasing function of i which means that $1+\sum_{j=1}^{N} P_{ij}[\sum_{m=1}^{M} S_{lm} V(j,m)]$ is also a non-increasing function of i while $(1-b_l)[1+\sum_{m'=1}^{M} S_{lm'} \sum_{r=1}^{N} R_r V(r,m')]$ is independent of i. Therefore, we can say that once we have a i^* where the inequality holds so that it is worth repairing, it will be worth repairing for $i \geq i^*$. Hence the optimal policy is repair if $i \geq i^*(m)$, do nothing if $i < i^*(m)$.

Theorem 3.2.3 $i^*(M) = N$.

(Proof)

This theorem holds if we can show

$$W_1(N-1, M) \geq W_2(M)$$

Since $S_{mm'}$ is stochastically ordered and $V(i, m)$ is monotonically non-increasing in m, if $l < M$,

$$\sum_{m'=1}^{M} S_{lm'} \sum_{r=1}^{N} R_r V(r, m') \geq \sum_{m'=1}^{M} S_{Mm'} \sum_{r=1}^{N} R_r V(r, m')$$

In (3.6), for $l < M$, since $(1 - b_l) \geq (1 - b_M)$,

$$\begin{aligned} W_2(l) &= (1 - b_l)[1 + \sum_{m'=1}^{M} S_{lm'} \sum_{r=1}^{N} R_r V(r, m')] \\ &\geq (1 - b_M)[1 + \sum_{m'=1}^{M} S_{Mm'} \sum_{r=1}^{N} R_r V(r, m')] = W_2(M) \end{aligned}$$

Hence,

$$W_2(l) \geq W_2(M)$$

Now, we can derive the following relation since $V(j, m') \geq W_2(m')$,

$$W_1(N-1, M) = 1 + \sum_{j=1}^{N} P_{N-1,j} \sum_{m'=1}^{M} S_{Mm'} V(N-1, m')$$

$$\geq \; 1 + \sum_{j=1}^{N} P_{N-1,j} \sum_{m'=1}^{M} S_{Mm'} W_2(M) = 1 + W_2(M) > W_2(M)$$

Therefore, $W_1(N-1, M) > W_2(M)$ which means that $i^*(M) = N$ in the most dangerous situation.

Lemma 3.2.3 $i^*(l) \leq i^*(M)$ for all l.

(Proof)

From theorem 3.2.1, we have already proved that $i^*(l) \leq N$ for any arbitrary environmental situation state l. Also, we have already proved $i^*(M) = N$. Hence, it is true that $i^*(l) \leq i^*(M)$ for all l.

Corollary 3.2.1 When $b_1 = b_2 = \cdots = b_M$, $i^*(1) = i^*(2) = \cdots = i^*(M) = N$

(Proof)

If $b_1 = b_2 = \cdots = b_M$, then by lemma 3.2.3 all environment states are equivalent to the M th state. By theorem 3.2.3, $i^*(M) = N$ and so hence by lemma 3.2.3, $i^*(1) = i^*(2) = \cdots = i^*(M) = N$.

3.3 Numerical Example

Description This numerical example has five different environment situations, 1(most peaceful environment), 2, 3, 4, 5 (most dangerous environment). We assume that there are total 10 unit quality states, 1(new), 2, · · ·, 9, 10 (down) with the Quality State Transition Probability Matrix(QSTPM) given by Table 3-1. We also assume that repair is not perfect but given by R_r in Table 3-3 which describes the probability that the quality state after the repair action is in state r. We show three numerical cases for this one repair action maintenance model. The probability of initiating event, b_m, where $1 \leq m \leq 5$ is for example 1 $(0.1, 0.2, 0.4, 0.6, 0.7)$ (i.e. gradual escalation), for example 2 $(0.1, 0.11, 0.4, 0.7, 0.9)$ (i.e. two good states and a gradual escalation), and for example 3 $(0.1, 0.11, 0.12, 0.15, 0.9)$ (i.e. four good states and one very dangerous state).

The results for example 1 and 2 are shown in Table 3-4 and Table 3-5. In example 1 since the probability of an initiating event in environment 2 is twice as big as that in environment 1, more repair actions occur in environment 1. In example 2, since the probability of initiating event in environment 3 is much higher than that of environment 1 and 2 which have a small difference between each other, it is best to do more repairs just before environment 3 is likely to occur. We can also find that the expected survival periods in example 2 are longer than those in example 1 for all states. This is because in example 2 there is more chance of surviving in environment 1 and 2 than in example 1.

The result for example 3 is shown in Table 3-6. In this case the probabilities of an initiating event in environmental situations 1 to 4 are very small and increase very slightly. However in situation 5 the probability of an initiating event increases dramatically from 0.15 to 0.9. The optimaility policy suggests repairing in more quality states when in environmental situation 4 than in the other situations including less dangerous situations. We can explain this because we are more likely to wish to repair before entering situation 5 where the probability of an initiating event increase so much.

Figures 3-9 and 3-10 show the overall trend of expected survival periods along the environment situation and quality state in each case. In these figures, only the integer points on the surface really exist. In state 10, one has to repair in both examples. In example 1, $i^*(1) = 7, i^*(2) = i^*(3) = i^*(4) = i^*(5) = 10$ and in example 2, $i^*(1) = 8, i^*(2) = 8, i^*(3) = i^*(4) = i^*(5) = 10$. For both examples, when the environmental situation is in 5, we repair the equipment only when it has failed.

Table 3-1. Quality State TPM

$i \setminus j$	1	2	3	4	5	6	7	8	9	10
1	0.2	0.2	0.2	0.1	0.08	0.05	0.05	0.05	0.05	0.02
2	0	0.2	0.2	0.2	0.1	0.1	0.08	0.05	0.04	0.03
3	0	0	0.2	0.2	0.2	0.1	0.1	0.1	0.05	0.05
4	0	0	0	0.2	0.2	0.2	0.15	0.1	0.1	0.05
5	0	0	0	0	0.2	0.3	0.2	0.1	0.1	0.1
6	0	0	0	0	0	0.2	0.3	0.2	0.2	0.1
7	0	0	0	0	0	0	0.2	0.3	0.3	0.2
8	0	0	0	0	0	0	0	0.3	0.4	0.3
9	0	0	0	0	0	0	0	0	0.4	0.6
10	0	0	0	0	0	0	0	0	0	1

Table 3-2. Environment Situation TPM

$m \setminus m'$	1	2	3	4	5
1	0.4	0.3	0.2	0.05	0.05
2	0.2	0.4	0.23	0.1	0.07
3	0.1	0.2	0.4	0.2	0.1
4	0.05	0.15	0.2	0.3	0.3
5	0.05	0.1	0.15	0.2	0.5

Table 3-3. Repair TPM

r	1	2	3	4	5	6	7	8	9	10
R_r	0.2	0.2	0.1	0.1	0.1	0.1	0.08	0.05	0.05	0.02

Table 3-4. Expected survival time under different actions and choice of optimal action for example 1

case	m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
1	1	0.1	W_1	14.2	13.9	13.6	13.4	13.1	12.9	12.6	12.4	12.1	10.7
			W_2	12.8	12.8	12.8	12.8	12.8	12.8	12.8	12.8	12.8	12.8
			Act.	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$
	2	0.2	W_1	14.0	13.8	13.4	13.2	12.9	12.6	12.2	11.9	11.4	8.9
			W_2	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	3	0.4	W_1	13.8	13.5	13.1	12.8	12.4	12.1	11.5	11.1	10.4	6.19
			W_2	8.61	8.61	8.61	8.61	8.61	8.61	8.61	8.61	8.61	8.61
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	4	0.6	W_1	13.5	13.2	12.7	12.4	12.0	11.5	10.8	10.3	9.3	3.9
			W_2	5.9	5.9	5.9	5.9	5.9	5.9	5.9	5.9	5.9	5.9
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	5	0.7	W_1	13.4	13.1	12.6	12.2	11.8	11.4	10.6	10.0	8.8	3.0
			W_2	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$

Table 3-5. Expected survival time under different actions and choice of optimal action for example 2

case	m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
2	1	0.1	W_1	17.1	16.8	16.5	16.3	16.0	15.8	15.4	15.2	14.7	12.9
			W_2	15.4	15.4	15.4	15.4	15.4	15.4	15.3	15.4	15.4	15.4
			Act.	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$
	2	0.11	W_1	17.0	16.7	16.4	16.2	15.9	15.6	15.3	15.0	14.5	12.5
			W_2	15.2	15.2	15.2	15.2	15.2	15.2	15.2	15.2	15.2	15.2
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$
	3	0.4	W_1	16.9	16.7	16.3	16.1	15.8	15.5	15.1	14.8	14.2	8.5
			W_2	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	4	0.7	W_1	16.0	15.6	15.1	14.6	14.1	13.5	12.5	11.7	10.2	3.3
			W_2	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	5	0.9	W_1	15.7	15.3	14.7	14.2	13.6	13.0	11.8	11.0	9.1	1.61
			W_2	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$

Table 3-6. Expected survival time under different actions and choice of optimal action for example 3

case	m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
3	1	0.1	W_1	19.7	19.5	19.2	19.0	18.8	18.5	18.3	18.1	17.9	15.9
			W_2	17.8	17.8	17.8	17.8	17.8	17.8	17.8	17.8	17.8	17.8
			Act.	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	0.11	W_1	19.6	19.3	19.1	18.8	18.6	18.3	18.0	17.8	17.5	15.4
			W_2	17.4	17.4	17.4	17.4	17.4	17.4	17.4	17.4	17.4	17.4
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	3	0.12	W_1	19.4	19.1	18.8	18.5	18.3	18.0	17.6	17.3	17.0	14.7
			W_2	17.1	17.1	17.1	17.1	17.1	17.1	17.1	17.1	17.1	17.1
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$	$\mathbb{R}$
	4	0.15	W_1	19.0	18.6	18.2	17.9	17.4	17.0	16.4	15.8	14.7	11.7
			W_2	16.2	16.2	16.2	16.2	16.2	16.2	16.2	16.2	16.2	16.2
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$
	5	0.9	W_1	18.7	18.3	17.8	17.4	16.8	16.4	15.4	14.6	12.8	2.0
			W_2	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8
			Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$

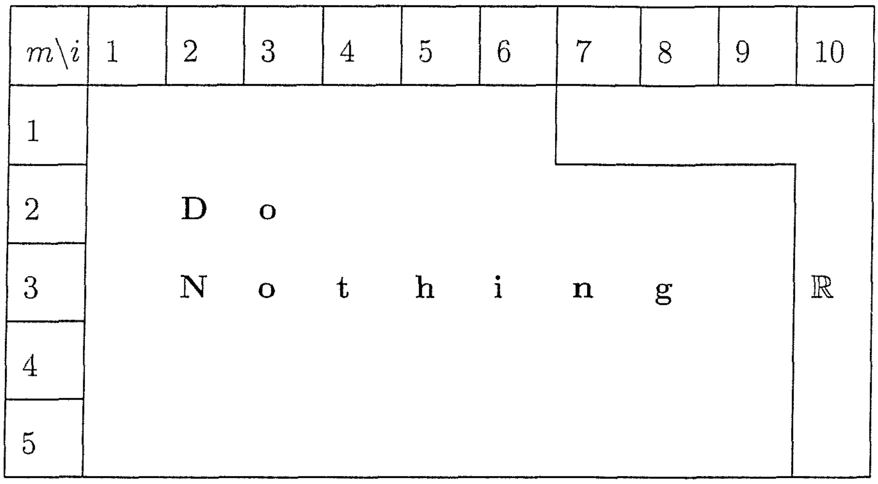

Figure 3-6 Result of example 1(simple form)

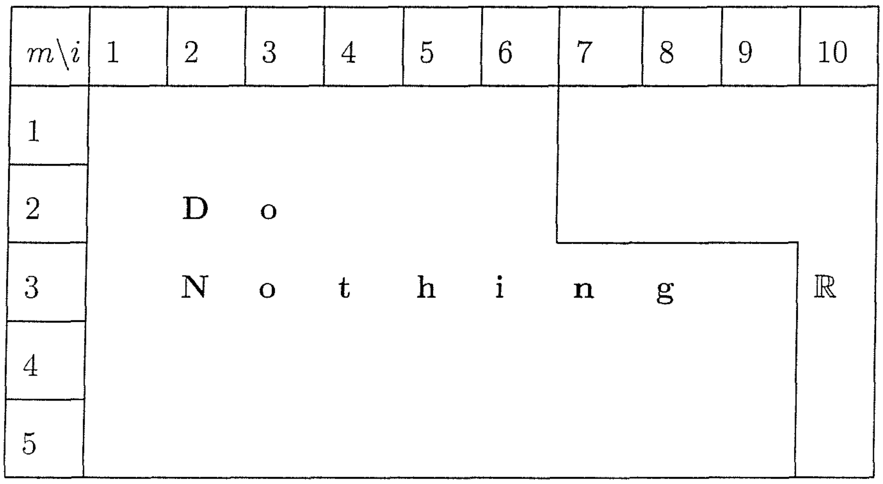

Figure 3-7 Result of example 2(simple form)

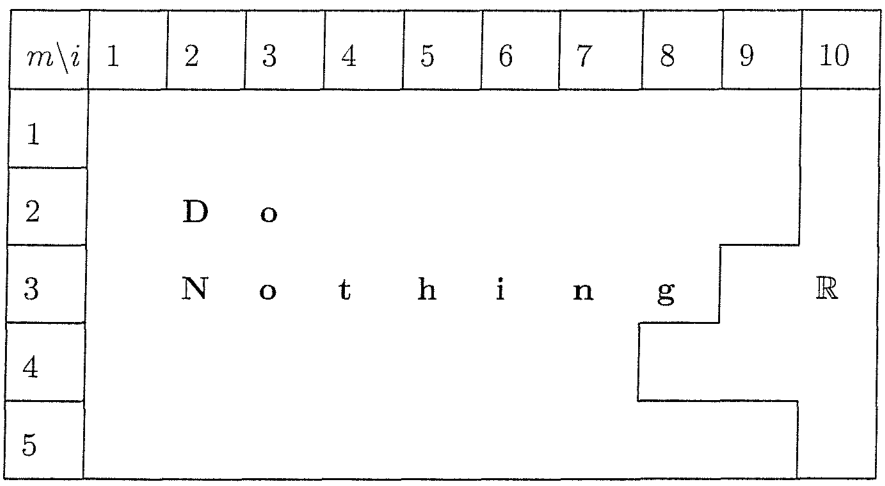

Figure 3-8 Result of example 3(simple form)

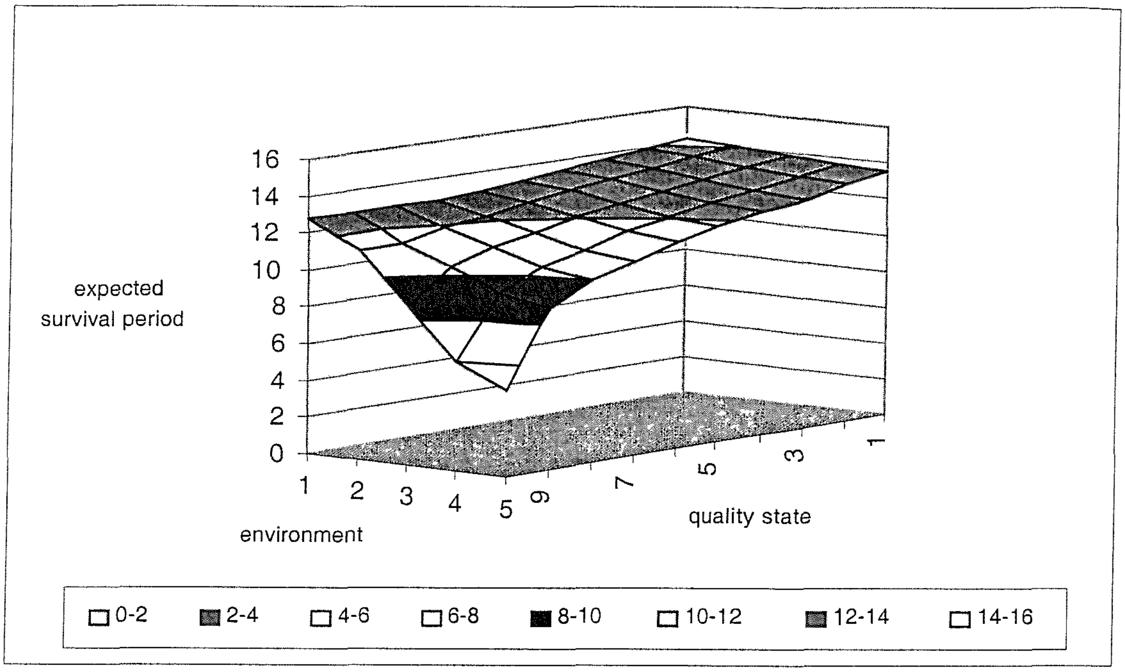

Figure 3-9. Expected Survival Period of example 1

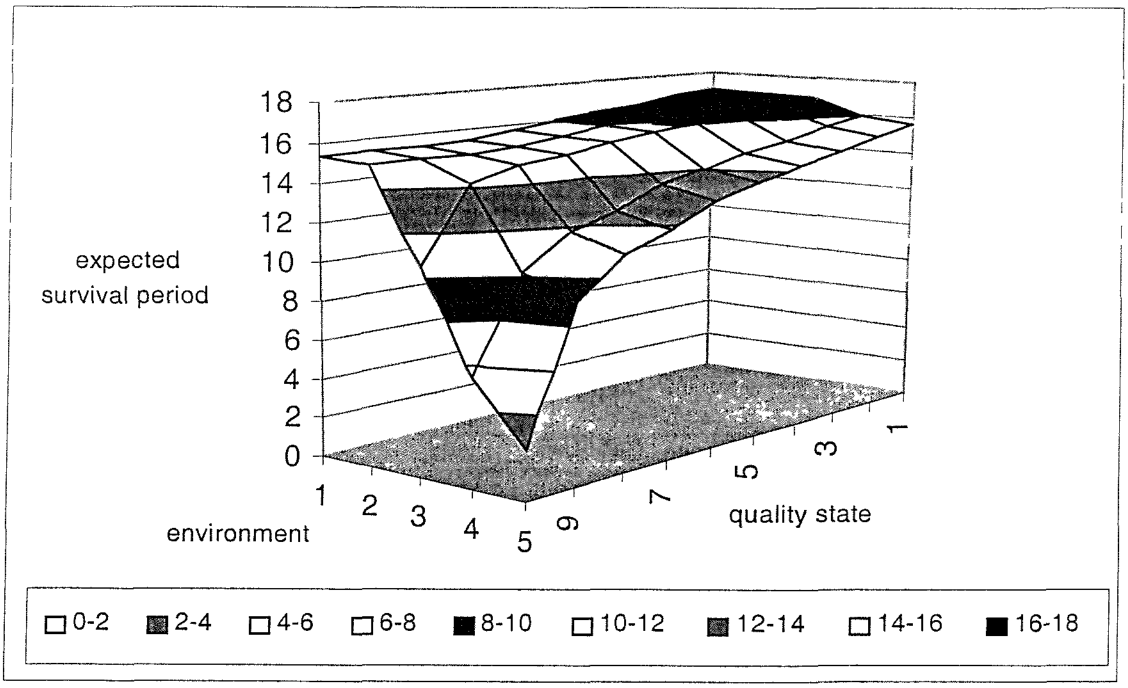

Figure 3-10. Expected Survival Period of example 2

3.4 Model for 2 Repair Actions under Multiple Environment

3.4.1 Introduction

In this section we consider two repair actions. One is the quick repair action and reflects repair in-situ while the other is a slow repair and reflects repair back at a central depot. The two repair actions differ both in time to repair and in the quality of the stand-by unit after the repair actions. Apart from the two different repair actions, all the other conditions are the same as those of the previous model.

3.4.2 Terminology

Possible Actions In this model, there are three possible maintenance actions by the operator at each time period.

(1) do nothing, (2) quick repair(low quality repair), (3) slow repair(high quality repair)

Repair The quick repair takes 1 unit time period until the completion of the repair. The slow repair takes 2 unit time periods until the completion of the repair. After repair, the standby unit has a different probability distribution R_r^q(for quick repair) and R_r^s(for slow repair) where

$$\sum_{r=1}^{N} R_r^q = 1, \sum_{r=1}^{N} R_r^s = 1, \sum_{r=1}^{k} R_r^q \leq \sum_{r=1}^{k} R_r^s (\text{stochastic ordering}) \text{ for all } k$$

3.4.3 Model

Model Description As in the 1 repair action model, we can consider several actions -do nothing and the two repair actions- for the each environment situation. The do nothing action can be divided into two sub-cases, (1) when the unit quality state is not in N(failure state), and (2) When the unit quality state is in N.

Case 1. Do Nothing
(1) $i \neq N$
(2) $i = N$
Case 2. QuickRepair
Case 3. Slow Repair

Hence, we need to develop four calculations for each environment situation, (1) a model for the maximum expected period with do nothing now($i \neq N$), $W_1(i,m)$, (2) a model for the maximum expected period with do nothing now($i = N$), $W_1(N,m)$, and (3) a model for the maximum expected period with quick repair now, $W_2(m)$. (4) a model for the maximum expected period with slow repair now, $W\ _3(m)$. Then, we can derive the optimality equation to decide which expected period among W_1, W_2 and W_3 is best.

Expected Survival Periods by Do Nothing under Environment Situation m, $W_1(i,m)$

The expected period by do nothing under the environment situation m is exactly the same as that of the 1 repair model.

(1) When the standby unit is in working condition $(i \neq N)$

$$W_1(i,m) = 1 + \sum_{j=1}^{N} P_{ij} \sum_{m'=1}^{M} S_{mm'} V_{n-1}(j,m')$$

(2) When the standby unit is in failure condition $(i = N)$

$$W_1(N,m) = (1 - b_m)[1 + \sum_{m'=1}^{M} S_{mm'} V_{n-1}(N,m')]$$

Expected Survival Periods by Quick Repair under Environment Situation m, $W_2(m)$

The quick repair action in this model is also the same as the repair action in the previous model.

$$W_2(m) = (1 - b_m)[1 + \sum_{m'=1}^{M} S_{mm'} \sum_{r=1}^{N} R_r^q V_{n-1}(r, m')]$$

Expected Survival Periods by Slow Repair under Environment Situation m, $W_3(m)$

If the standby unit is in slow repairing which takes 2 time unit periods, the vulnerable period is different. If there is an occurrence of an initiating event, the system has a catastrophic event automatically. For this case we can consider the following three possible situations: (1) an initiating event occurring in the first repair period, (2) no initiating event in the first repair period, but an initiating event occurring in the second repair period, (3) no initiating event occurring during the repair period. We show these possible cases from the Figure 3-11, 3-12, and 3-13.

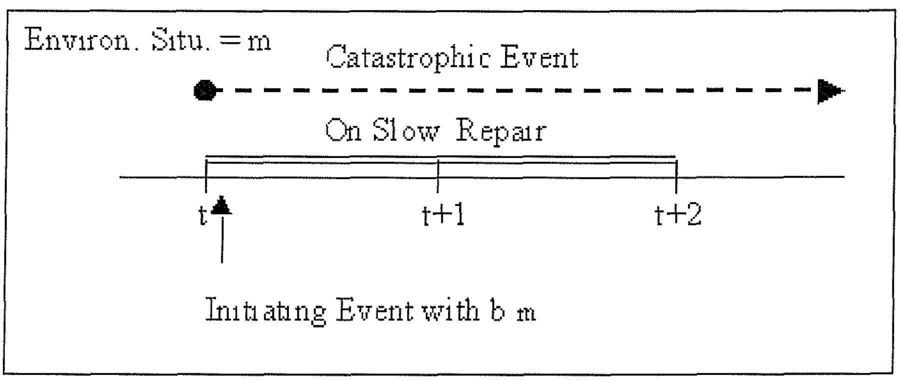

Figure 3-11. An initiating event occuring in the first repair period

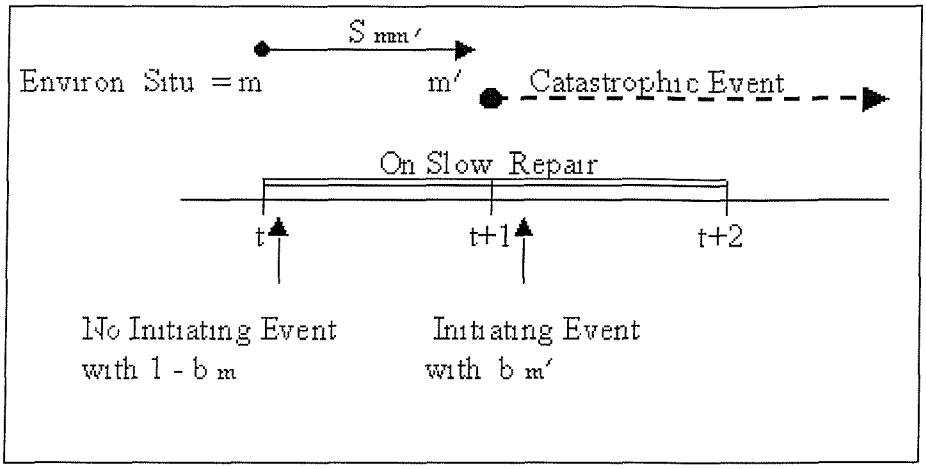

Figure 3-12. An initiating event occuring in the second repair period

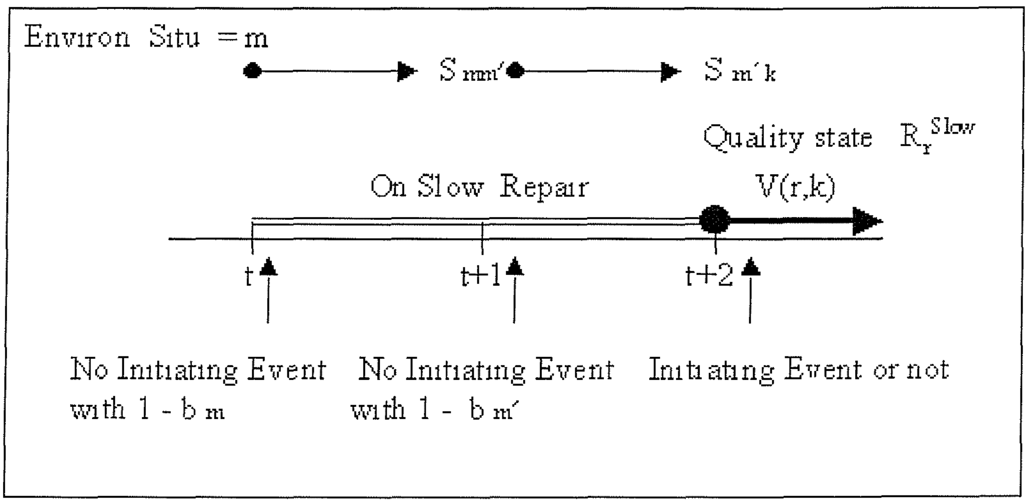

Figure 3-13. No initiating event occuring during repair period

Hence, the expected survival period by slow repair when the environment situation is m, $W_3(m)$ is

$$W_3(m) = (1 - b_m)[2 - \sum_{m'=1}^{M} S_{mm'} b_{m'} + \sum_{m'=1}^{M} S_{mm'}(1 - b_{m'}) \sum_{k=1}^{M} S_{m'k} \sum_{r=1}^{N} R_r^s V(r,k)]$$

Optimality Equation We derive the optimality equation to select the best repair action in each state. In order to do this, we compare the expected periods until a catastrophic event if we do nothing, enact a quick repair or a slow repair at this period. The optimality equation is:

$$V(i,m) = \max\{W_1(i,m), W_2(m), W_3(m)\} \tag{3.7}$$

,where

$$W_1(i,m) = (1-\delta_{iN}b_m)[1+\sum_{j=1}^{N} P_{ij}\sum_{m'=1}^{M} S_{mm'}V(j,m')] \tag{3.8}$$

where

$$\delta_{iN} = 1 \text{ if } i = N$$

$$\delta_{iN} = 0 \text{ otherwise}$$

$$W_2(m) = (1-b_m)[1+\sum_{m'=1}^{M} S_{mm'}\sum_{r=1}^{N} R_r^q V(r,m')] \tag{3.9}$$

$$W_3(m) = (1-b_m)[2-\sum_{m'=1}^{M} S_{mm'}b_{m'} + \sum_{m'=1}^{M} S_{mm'}(1-b_{m'})\sum_{k=1}^{M} S_{m'k}\sum_{r=1}^{N} R_r^s V(r,k)] \tag{3.10}$$

$(3.7), (3.8), (3.9), (3.10)$ can be solved using value iteration. The results of the previous section extend to this model and the proofs follow by induction on value iteration. The value iteration scheme satisfies equation $(3.7), (3.8), (3.9), (3.10)$ with V_n, W_n on the L.H.S. and V_{n-1} on the R.H.S.

Lemma 3.4.1 $V(i,m)$ is a non-increasing function of i, where i is the quality state, $1 \leq i \leq N$, and m is the arbitrary environmental situation state, $1 \leq m \leq M$.

(Proof)

The proof is given by induction on n in $V_n(\cdot,\cdot)$ where $V_n(\cdot,\cdot)$ is the nth iterate of value iteration. Using the induction hypothesis, the stochastic ordering of P_{ij} in i and $V_n(i,m)$ being a non-increasing function of i, we get the following result for (3.8):

$$W_n^1(i,m) > W_n^1(i+1,m)$$

$W_n^2(m)$ and $W_n^3(m)$ for i and $i+1$ are equal. Therefore, we can conclude that $V_n(i,m) \geq V_n(i+1,m)$ which means that $V_n(i,m)$ is a non-increasing function of i. Hence in the limit as $n \to \infty$, $V(i,m)$ is a non-increasing function of i.

Lemma 3.4.2 $V(i,l) \geq V(i,l')$ for all i, where $l \leq l'$

(Proof)

The proof follows as in Lemma 3.4.1

Theorem 3.4.1 In state N(down), one should always repair.

(Proof)

From Theorem 3.2.1, we know if there was only a quick repair action one would use the quick repair in state N. Adding a slow repair cannot decrease the state where repair occurs and here the result holds.

Theorem 3.4.2 $i^*(M) = N$.

(Proof)

From the previous theorem, we proved that at state N(down) we always need to repair. This theorem requires one to prove

$$W_1(N-1, M) \geq \max[W_2(M), W_3(M)]$$

The proof that $W_1(N-1, M) > W_2(M)$ follows exactly as in Theorem 3.2.3. Now we need to prove $W_1(N-1, M) > W_3(M)$,

$$\begin{aligned} W_3(l) &= (1-b_l)[2 - \sum_{m'=1}^{M} S_{lm'} b_{m'} + \sum_{m'=1}^{M} S_{lm'}(1-b_{m'}) \sum_{k=1}^{M} S_{m'k} \sum_{r=1}^{N} R_r^s V(r,k)] \\ W_3(M) &= (1-b_M)[2 - \sum_{m'=1}^{M} S_{Mm'} b_{m'} + \sum_{m'=1}^{M} S_{Mm'}(1-b_{m'}) \sum_{k=1}^{M} S_{m'k} \sum_{r=1}^{N} R_r^s V(r,k)] \end{aligned}$$

Using the result of Theorem 3.2.1, we get

$$W_1(N-1, M) \geq 1 + \sum_{m'=1}^{M} S_{Mm'} W_3(m^{'})$$

Since $b_l \leq b_M$, $\sum_{m'=1}^{M} S_{lm'} b_{m'} \leq \sum_{m'=1}^{M} S_{Mm'} b_{m'}$ (stochastic ordering), $\sum_{m'=1}^{M} S_{lm'}(1-b_{m'}) \geq \sum_{m'=1}^{M} S_{Mm'}(1-b_{m'})$ (stochastic ordering), we can say

$$\begin{aligned} & W_3(l) \\ = \; & (1-b_l)[2 - \sum_{m'=1}^{M} S_{lm'} b_{m'} + \sum_{m'=1}^{M} S_{lm'}(1-b_{m'}) \sum_{k=1}^{M} S_{m'k} \sum_{r=1}^{N} R_r^s V(r,k)] \\ \geq \; & (1-b_M)[2 - \sum_{m'=1}^{M} S_{Mm'} b_{m'} + \sum_{m'=1}^{M} S_{Mm'}(1-b_{m'}) \sum_{k=1}^{M} S_{m'k} \sum_{r=1}^{N} R_r^s V(r,k)] \\ = \; & W_3(M) \end{aligned}$$

Hence, we can confirm that

$$
\begin{aligned}
& W_1(N-1,M) \\
\geq\ & 1+\sum_{m'=1}^{M} S_{Mm'}W_3(m') \geq 1+\sum_{m'=1}^{M} S_{Mm'}W_3(M) \\
=\ & 1+W_3(M) \geq W_3(M)
\end{aligned}
$$

Thus $W_1(N-1, M) > \max[W_2(M), W_3(M)]$ which means we always have $i^*(M) = N$.

<u>Lemma 3.4.3</u> $i^*(l) \leq i^*(M)$ for all l.

(Proof)

From theorem 3.4.1, we proved that $i^*(l) \leq N$ for an arbitrary environmental situation l. Also, from theorem 3.4.2, we know that $i^*(M) = N$. Hence, it is true that $i^*(l) \leq i^*(M)$ for all i.

<u>Corollary 3.4.1</u> When $b_1 = b_2 = ---- = b_M = b$, $i^*(1) = i^*(2) = ---- = i^*(M) = N$

(Proof)

If $b_1 = b_l = b_M$, by lemma 3.4.3 all environmental state are equivalent to state M. By theorem 3.4.2, $i^*(M) = N$. And so by lemma 3.4.3, $i^*(1) = i^*(2) = --- = i^*(M) = N$.

3.5 Numerical Example

<u>Description</u> In this example, there are also 5 different environment situation states and10 different unit quality states. The probability of an initiating event,

b_m, where $1 \leq m \leq 5$ is $(0.1, 0.2, 0.4, 0.6, 0.7)$. The transition matrices for both quality state and environment situation are the same as in example 3. The effect of the repairs has the following distributions. These are different from the repair distribution in example 1 to 3 in that the quick repair tends to return the item to a worse state and the slow repair tends to return the item to a better state than the repair action in those examples.

Table 3-7. Repair TPM for example 4

State	1	2	3	4	5	6	7	8	9	10
Quick Repair, R_r^q	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.2	0.2	0.4
Slow Repair, R_r^s	0.8	0.1	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01

The results in Table 3-8 show the results are similar to those in the previous numerical examples. This model recommends repairing the stand-by unit before environment 2 because the initiating event probability, $b_2 = 0.2$ is twice as big as $b_1 = 0.1$. However the repair is not the quick repair but the slow repair. In environmental situations 2 and 3 , we also need to do the slow repair action at quality state 10. The reason is that the probability of an initiating event in this situation is much lower than that in situations 4 and 5. Hence we prefer a slow repair because of the better repair quality despite the fact that it takes twice as long as the quick repair action. In situation 4 and 5 we need to do a quick repair because the initiating event probability is so much higher.

In example 5, the probability of an initiating event, b_m, where $1 \leq m \leq 5$ is $(0.1, 0.11, 0.12, 0.15, 0.9)$. The quality after repair follows the distributions in

Table 3-8. The Table 3-9 shows the environment situation transition probabilities for each environment situation.

Table 3-8. Repair TPM for example 5

State	1	2	3	4	5	6	7	8	9	10
Quick Repair, R_r^q	0.0	0.0	0.0	0.0	0.2	0.3	0.2	0.1	0.1	0.1
Slow Repair, R_r^s	0.8	0.1	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table 3-9. Environment Situation TPM for example 5

$m \setminus m'$	1	2	3	4	5
1	0.9	0.05	0.03	0.01	0.01
2	0.6	0.2	0.1	0.06	0.04
3	0.38	0.28	0.18	0.07	0.09
4	0.04	0.06	0.1	0.2	0.6
5	0.01	0.01	0.03	0.05	0.9

As in the result of example 5, we have an example where quick repair is used and even one where quick repair can occur for an up state in one environmental situation, but one does nothing in better and worse environmental situations.

Table 3-10. Expected survival time under different actions and choice of optimal action for example 4

m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
1	0.1	W_1	10.0	9.8	9.5	9.2	9.0	8.7	8.3	8.1	7.7	6.7
		W_2	7.4	7.4	7.4	7.4	7.4	7.4	7.4	7.4	7.4	7.4
		W_3	8.2	8.2	8.2	8.2	8.2	8.2	8.2	8.2	8.2	8.2
		Act.	DN	DN	DN	DN	DN	DN	DN	$\mathbb{SR}$	$\mathbb{SR}$	$\mathbb{SR}$

m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
2	0.2	W_1	9.9	9.7	9.3	9.1	8.8	8.4	8.0	7.7	7.2	5.4
		W_2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2
		W_3	6.8	6.8	6.8	6.8	6.8	6.8	6.8	6.8	6.8	6.8
		Act.	DN	DN	DN	DN	DN	DN	DN	DN	DN	$\mathbb{SR}$

m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
3	0.4	W_1	9.8	9.5	9.1	8.8	8.5	8.1	7.5	7.1	6.4	3.4
		W_2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2	4.2
		W_3	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4
		Act.	DN	DN	DN	DN	DN	DN	DN	DN	DN	$\mathbb{SR}$

m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
4	0.6	W_1	9.6	9.3	8.9	8.5	8.1	7.7	7.1	6.6	5.6	1.9
		W_2	2.6	2.6	2.6	2.6	2.6	2.6	2.6	2.6	2.6	2.6
		W_3	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4
		Act.	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	**QR**

m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
5	0.7	W_1	9.5	9.2	8.7	8.4	8.0	7.6	6.8	6.3	5.3	1.3
		W_2	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9
		W_3	1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.7
		Act.	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	**QR**

$m \backslash i$	1	2	3	4	5	6	7	8	9	10
1										$\mathbb{SR}$
2		**D**	**o**							
3		**N**	**o**	**t**	**h**	**i**	**n**	**g**		
4										**QR**
5										

Figure 3-14 Result of example 4(simple form)

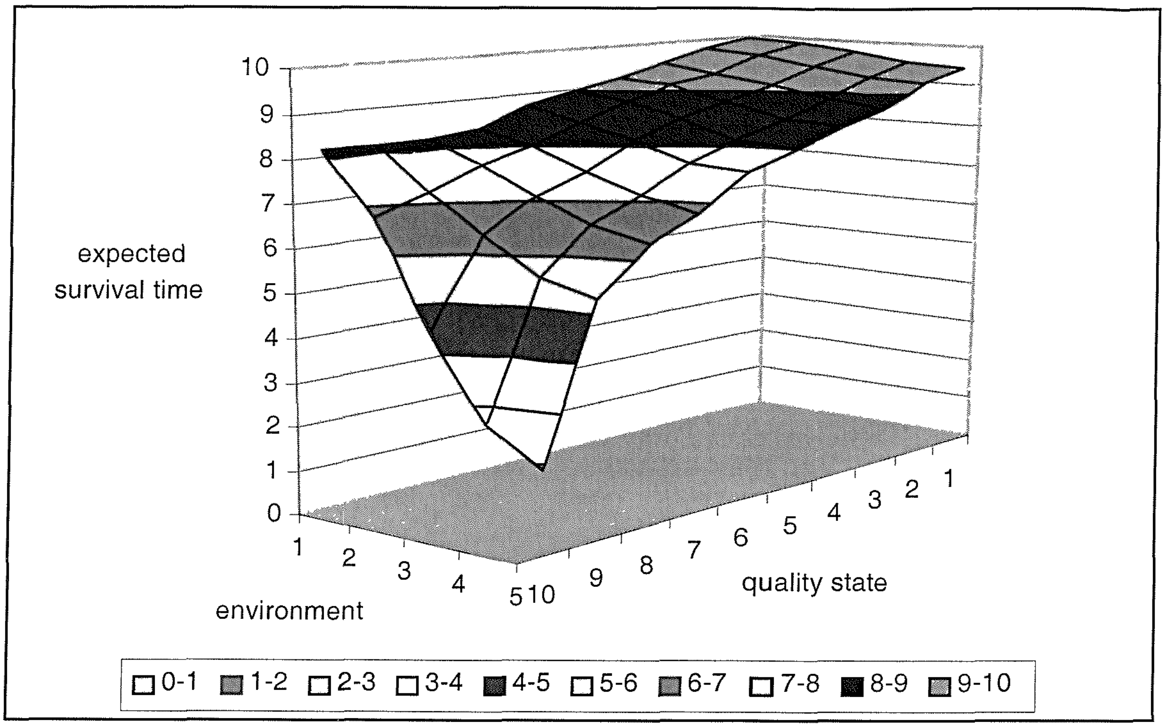

Figure 3-15. Expected Survival Period in example 4

Table 3-11. Expected survival time under different actions and choice of optimal action for example 5

m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
1	0.1	W_1	22.1	21.9	21.6	21.4	21.2	21.0	20.7	20.5	20.0	17.7
		W_2	19.1	19.1	19.1	19.1	19.1	19.1	19.1	19.1	19.1	19.1
		W_3	18.8	18.8	18.8	18.8	18.8	18.8	18.8	18.8	18.8	18.8
		Act.	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	*DN*	**QR**

m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
2	0.2	W_1	20.8	20.5	20.3	20.1	19.8	19.6	19.2	19.0	18.6	16.2
		W_2	17.6	17.6	17.6	17.6	17.6	17.6	17.6	17.6	17.6	17.6
		W_3	17.3	17.3	17.3	17.3	17.3	17.3	17.3	17.3	17.3	17.3
		Act.	DN	DN	DN	DN	DN	DN	DN	DN	DN	**QR**

m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
3	0.4	W_1	19.7	19.5	19.2	18.9	18.7	18.4	18.0	17.7	17.3	14.8
		W_2	16.4	16.4	16.4	16.4	16.4	16.4	16.4	16.4	16.4	16.4
		W_3	16.0	16.0	16.0	16.0	16.0	16.0	16.0	16.0	16.0	16.0
		Act.	DN	DN	DN	DN	DN	DN	DN	DN	DN	**QR**

m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
4	0.6	W_1	12.1	11.7	11.2	10.7	10.3	9.7	9.0	8.5	7.6	5.6
		W_2	8.7	8.7	8.7	8.7	8.7	8.7	8.7	8.7	8.7	8.7
		W_3	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5
		Act.	DN	DN	DN	DN	DN	DN	DN	**QR**	**QR**	**QR**

m	$b_m \backslash i$		1	2	3	4	5	6	7	8	9	10
5	0.7	W_1	10.0	9.6	8.9	8.4	7.8	7.2	6.3	5.6	4.2	0.3
		W_2	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
		W_3	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
		Act.	DN	DN	DN	DN	DN	DN	DN	DN	DN	**QR**

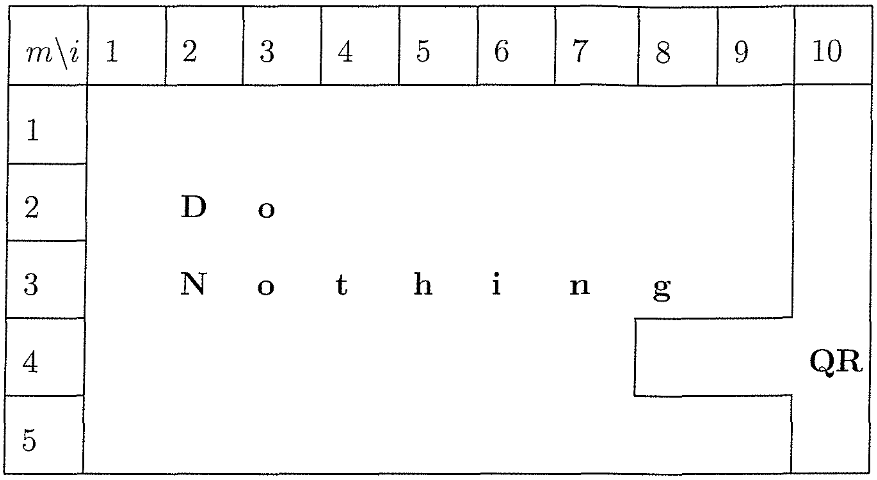

Figure 3-16 Result of example 5(simple form)

3.6 Conclusions

The models presented in this paper show that there is a strong interaction between the quality state of the stand-by unit, the general environment state, and the repair action. In the one repair action case,one always repairs when the unit is down and the most hostile environment only repairs when the unit is down. In other environments one can make repairs when the unit is still functioning. We have produced a counter example to the assumption that the states at which one repairs increase as the environment becomes more benign.

With quick and slow repairs, the repair or do nothing decision continue to have the same proportions as in the one action case. The choice between quick or slow repair depends critically on the relative outcomes of the two repair processes. The numerical example suggests one is more likely to move from slow to quick repairs as the environment becomes more hostile, but this need not always be the case.We have examples where quick repair(or slow repair) is used and even one where quick repair can occur for an up state in one environmental situation(see Figure 3-16 of example 5), but one does nothing in better and worse environmental situations.

CHAPTER 4

Repair Strategies in an Uncertain Environment: Stochastic Game Approach

4.1 Introduction

In chapter 3, we have developed discrete time Markov decision process formulations of the standby unit repair problem in order to investigate the form of the optimal repair policies which maximise the expected survival time until a catastrophic event. In chapter 4, we look at conflict situations where the environment is controlled by an opponent. In this case the opponent's actions force the need for the equipment, and this situation is modelled as a stochastic game.

For this, we develop stochastic game models in global and local constraints on effort. In the models with global constraints on effort, we consider all time periods and give a capacity limit to the player II. For the local case we only need to consider the opponent's action over a short previous time horizon. In our stochastic game model, we discretize the state space. For the decayed capacity limit model of this stochastic game model, we use approximation methods in order to get a sleep index. Using this method we can make the continuous value of the sleep index to the discrete one. Also, for the pure capacity limit model, in which there is no discounted factor, we use boundary condition because the sleep index goes to infinity. For we do this by assuming that if the sleep index is bigger than a certain number N, we assume that the index is equal to N. The outline of this chapter is as follows.

In section 4.2, we look at the problem where there are no constraints on the enemy in terms of the actions they can perform. This leads to a complete but unrealistic solution to the problem because in reality, an opponent- be it rogue country or terrorist - cannot be continually on the attack. Thus in section 4.3 we introduce the idea of a constraint on the average effort undertaken by the opponent over the total history of the game so far. We naively describe this as a sleep index in that the opponent needs to sleep for a certain percentage of the time. This reflects the need to be able to regroup, resupply and replan between times of high levels of actions. In section 4.3 we further expand these results to the situation where the advantage of a rest or quiescent period is discounted the further in the past it is, but always has a positive effect. In section 4.5 we look at games where the benefit to the opponent of being " able to sleep" only lasts for a finite period and is then lost completely. In each case we are able to derive properties of the form of the optimal maintenance policy and also to find the form of the optimal policy in specific numerical examples. In the concluding section we consider the implication of these results.

4.2 Basic Stochastic Game Model

4.2.1 Introduction Consider the basic game where player I has to decide when to undertake preventive maintenance or repair on his stand-by equipment and player II has to decide what level of threat should be in the environment. In the military context, player I is the friendly forces and player II the enemy forces. The game is played repeatedly and at each period, player I has to decide whether to perform maintenance/repair on the equipment or to do nothing. We assume the environment can be in one of two states, - state 1, which is a "peaceful" environment and state 2 , which is a more "dangerous" environment- and player II chooses at each period what the environment state will be. These are akin to the DEFCON states defined by U.S. and other military authorities for assessing the environment. The chance of an initiating event, which requires the stand-by system to respond is b_m if the environment in state m, with $b_1 \leq b_2$. The stand-by unit is inspected regularly each period and this gives information on the operational state of the equipment. These states can represent either operational states, such as A, B, C, D used by same military or could be functions of age and operational history. Assume the equipment can be in one of N states, ranging from 1, as good as new, via $2, \cdots, N-1$ which are still conditions where it is still operable to state N which means the equipment is down and will not work. If no action is performed on the equipment it will move from state i at one period to state j at the next with probability P_{ij} where $\sum_{j=1}^{N} P_{ij} = 1$ and $P_{NN} = 1$. The "ordering" of the intermediate equipment states reflects increasing pessimism about their future operability and this is expressed by assuming P_{ij} satisfies a first order stochastic ordering condition namely

$$\sum_{j<k} P_{ij} \geq \sum_{j<k} P_{i+1,j} \text{ for all } i = 1, \cdots, N-1, k = 1, \cdots, N,$$

If the preventive maintenance/repair action is performed (the former if equipment state is $i = 1, \cdots, N-1$, the latter if the state is N), this takes one time period, during which the equipment cannot be used, if required, and returns the equipment to the state1 - the good as new state. The subsequent results also hold if the maintenance action is not perfect, and returns the equipment to state i with probability R_i, but we will not complicate notation by describing this case.

<u>4.2.2 Game Model</u> Player I's aim is to maximise the time until a catastrophic event occurs- an initiating event is triggered and the stand-by system is unable to respond. Player II on the other hand wants to minimise the time until the catastrophic event occurs and knows the state of I's equipment. Thus this situation where the two players are completely at odds with one another can be modelled as a two players zero sum stochastic game as follows.

Let the game Γ have N subgames $\Gamma_1, \cdots, \Gamma_N$ where Γ_i corresponds to the equipment being in state i. Player I then decides whether to perform a maintenance action or do nothing for the next period while at the same time player II decides whether to act so the environment is peaceful(state1) or dangerous(state 2). This defines the probability that an initiating event will occur during the period, and hence if the equipment is down or being repaired, whether there is a catastrophic event. If the equipment is in state $i(\Gamma_i)$ and no maintenance is carried out, it will move to state j(and the game to subgame Γ_j) for the next period with probability P_{ij}. Thus the payoff matrix when the game is in subgame Γ_i is given by

$\Gamma_i, i \neq N$		II	
		Making Peaceful Situation	Making Dangerous Situation
I	Do Nothing	$1 + \sum_{j=1}^{N} P_{ij}\Gamma_j$	$1 + \sum_{j=1}^{N} P_{ij}\Gamma_j$
	Repair	$(1 - b_1)(1 + \Gamma_1)$	$(1 - b_2)(1 + \Gamma_1)$

(4.1)

Γ_N		II	
		Making Peaceful Situation	Making Dangerous Situation
I	Do Nothing	$(1 - b_1)(1 + \Gamma_N)$	$(1 - b_2)(1 + \Gamma_N)$
	Repair	$(1 - b_1)(1 + \Gamma_1)$	$(1 - b_2)(1 + \Gamma_1)$

(4.2)

In (4.1), the numbers represent the reward in this stage of the game and the $P_{ij}\Gamma_j$ implies that with probability P_{ij} the next play of the game will be subgame Γ_j. For discounted stochastic games, which these are because of the $(1 - b_i)$ term which guarantees the value of the game is bounded above by $\frac{1}{1-P_{11}} \times \frac{1}{b_1}$, the value of the game, v_i, starting in subgame i, for $i = 1, \cdots, N$ satisfies the equation

$$v(i) = val \begin{bmatrix} 1 + \sum_{j=1}^{N} P_{ij}v(j) & 1 + \sum_{j=1}^{N} P_{ij}v(j) \\ (1 - b_1)(1 + v(1)) & (1 - b_2)(1 + v(1)) \end{bmatrix}, i \neq N \tag{4.3}$$

or

$$v(N) = val \begin{bmatrix} (1-b_1)(1+v(N)) & (1-b_2)(1+v(N)) \\ (1-b_1)(1+v(1)) & (1-b_2)(1+v(1)) \end{bmatrix} \tag{4.4}$$

Moreover the solution of this game with an infinite number of periods can be solved using a value iteration approach where the n^{th} iterate $v_n(i)$ (which corresponds to value if only n periods allowed) satisfies

$$v_n(i) = val \begin{bmatrix} 1+\sum_{j=1}^{N} P_{ij}v_{n-1}(j) & 1+\sum_{j=1}^{N} P_{ij}v_{n-1}(j) \\ (1-b_1)(1+v_{n-1}(1)) & (1-b_2)(1+v_{n-1}(1)) \end{bmatrix}, i \neq N \tag{4.5}$$

with a similar equation based on (4.4) for $v_n(N)$. $v_0(i) = 0$ by definition. This allows us to solve the game with help of the following lemmas.

<u>Lemma 4.2.1</u> $v_n(i)$ is non-deceasing in n and non-increasing in i and converges to $v(i)$.

<u>Proof</u> The non-decreasing result in n follows since $v_1(i) \geq v_0(i) = 0$ and then by induction. Since $v_{n-1}(i) \geq v_{n-2}(i)$ for all i the four terms in the payoff matrix for $v_n(i)$ are greater than or equal to the four terms in the matrix for $v_{n-1}(i)$. Hence $v_n(i) \geq v_{n-1}(i)$ and the induction step is proved.

Similarly $0 = v_0(i+1) \leq v_0(i) = 0$ for all i, so the hypothesis holds for $n = 0$. Assume true for $v_{n-1}(i)$ then the stochastic ordering plus the monotonicity of $v_{n-1}(i)$ implies $\sum_{j=1}^{N} P_{i+1,j}v_{n-1}(j) \leq \sum_{j=1}^{N} P_{ij}v_{n-1}(j)$. Hence again each of the four entries in (4.5) of $v_n(i)$ is as large if not larger than the corresponding terms for $v_n(i+1)$, so $v_n(i+1) \leq v_n(i)$ for $i = 1, \cdots, N-1$. The same result holds for

$v_n(N) \leq v_n(N-1)$ since for $v_n(N)$ it is clear that repair dominates do nothing because $v_{n-1}(N) \leq v_{n-1}(1)$. Hence $v_n(N) = \min\{(1-b_1)(1+v_{n-1}(1)), (1-b_2)(1+v_{n-1}(1))\} \leq v_n(N-1)$ and the induction step holds.

Trivially since $v_n(i) \leq v_{n+1}(i)$, $v_n(\cdot)$ converges to $v(\cdot)$ because $v_n(i)$ is a bounded increasing function, bounded by T/b_1 where T is expected number of periods from state 1 to state N under the do nothing transition matrices. Similarly since $v_n(i) \geq v_n(i+1)$ and $v_n(\cdot)$ converges to $v(\cdot)$, we get the following corollary.

Corollary 4.2.1 $v(i)$ is non-increasing in i.

Theorem 4.2.1 The optimal strategy in the unconstrained game is : for player II always to choose the dangerous environment, for player I to do nothing in states $i < i^*$, where $i^* \leq N$, and perform maintenance/repair in state i^* to N.

Proof Since $b_2 > b_1$, trivially the dangerous strategy for player II always dominates the peaceful strategy. Since by the corollary $v(1) \geq v(N)$, the repair strategy(for the dangerous environment) is as good if not better than do nothing strategy for state N. The monotonicity of $v(j)$ together with the stochastic ordering of P_{ij} implies $\sum P_{ij}v(j)$ is non-increasing in i and so once it goes below $(1-b_2)(1+v(1))$ (: the definition of i^*) it will remain below it for all higher states i.

The solution of this game is unrealistic because the assumption that player II can always ensure the environment is dangerous is unrealistic. In the next section we remove this assumption.

4.3 Model with Global Constraints on Effort

4.3.1 Introduction One reason an enemy cannot continuously create a dangerous environment, is that it needs time to regroup, plan and rest its forces- which we facetiously describe as "sleep". One possible assumption is that at any stage n of the game, the enemy can only have created a dangerous environment in at most a proportion c of the previous stages. Thus if it has created a dangerous situation in d stages, $d \leq cn$ and $s = cn - d$ is a measure of the "sleep index", how much effort the enemy still has available to create dangerous situations. If the sleep index is s and at the next period player II chooses a peaceful environment, the index will move to $s + c$, while if he chooses a dangerous environment, the index will move to $s - (1 - c) = s + c - 1$. In this model the effect of the rest induced by a peaceful environment will endure undiminished throughout all the future. An alternative view is that the c value that the restful period adds to the "sleep" index should diminish to αc next period $\alpha^2 c$ the period thereafter and so on. This is equivalent to saying that if the current sleeping index is s, and this period player II keeps the environment peaceful, the index will move to $\alpha s + c$, while if player II chooses to make the environment dangerous the index will move to $\alpha s - (1 - c)$.

4.3.2 Game Model We will prove results for the two cases $\alpha = 1$(undiscounted) and $\alpha < 1$(discounting of the index) in the same model though in the former case the sleep index could be infinite, while in the latter case it is bounded above by $c/(1 - \alpha)$. In order always to have a finite set of subgames, we will always assume in the undiscounted case that the index cannot exceed M. So the stochastic game Γ model of this situation consists of a series of subgame $\Gamma_{i,s}$, where $i = 1, \cdots, N$ and $0 \le s \le \min\{M, c/(1-\alpha)\}$. Although the sleep index set appears continuous, it is in fact countable infinite, and in fact finite if only r stages are allowed. If the index starts with s_0 then after r stages, the value can only be $\alpha^r s_0 + c(1-\alpha^r)/(1-\alpha) - \sum_{i=1}^{r} Z_i \alpha^{r-i}$ where $Z_i = 1$ or 0 depending on where player II played dangerous or peaceful at the i^{th} stage.

Let $v(i, s)$ be the value of the game starting in $\Gamma_{i,s}$, then the values satisfy the equations

$$v(i,s) = val \begin{bmatrix} a_{11}, & a_{12} \\ a_{21}, & a_{22} \end{bmatrix}$$

where

$$\begin{aligned} a_{11} &= (1 - \delta_N(i) b_1)(1 + \sum_{j=1}^{N} P_{ij} v(j, \alpha s + c) \\ a_{12} &= (1 - \delta_N(i) b_2)(1 + \sum_{j=1}^{N} P_{ij} v(j, \alpha s + c - 1)) \\ a_{21} &= (1 - b_1)(1 + v(1, \alpha s + c)) \\ a_{22} &= (1 - b_2)(1 + v(1, \alpha s + c - 1)) \end{aligned}$$

and

$$\delta_N(i) = 1 \text{ if } i = N, \ 0 \text{ otherwise}$$

The value iteration algorithm defines $v_n(i)$ by

$$v_n(i,s) = val \begin{bmatrix} a^n_{11}, & a^n_{12} \\ a^n_{21}, & a^n_{22} \end{bmatrix} \tag{4.6}$$

where

$$\begin{aligned}
a_{11}^n &= (1-\delta_N(i)b_1)(1+\sum_{j=1}^{N} P_{ij}v_{n-1}(j,\alpha s+c) \\
a_{12}^n &= (1-\delta_N(i)b_2)(1+\sum_{j=1}^{N} P_{ij}v_{n-1}(j,\alpha s+c-1)) \\
a_{21}^n &= (1-b_1)(1+v_{n-1}(1,\alpha s+c)) \\
a_{22}^n &= (1-b_2)(1+v_{n-1}(1,\alpha s+c-1))
\end{aligned}$$

where again if $\alpha s+c \geq M$, it will be defined as M. Again define $v_0(i,s)=0$ for all i and s. As in section two, in order to prove results about the optimal policies for the game, Γ, one proves results about $v_n(i,s)$ and hence $v(i,s)$.

Lemma 4.3.1 $v_n(i,s)$ is non-deceasing in n and non-increasing in i and s.

Proof All the results follow from induction and the fact that if $W_1 = val \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}$ and $W_2 = val \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}$ then if $a_1 \geq a_2, b_1 \geq b_2, c_1 \geq c_2, d_1 \geq d_2, W_1 \geq W_2$. Since $v_0(i,s)=0, v_1(i,s)\geq 0, v_1(i,s)-v_0(i,s)\geq 0$. Assume that for $v_n(i,s)\geq v_{n-1}(i,s)$ for all i,s and note that from (4.6) $v_{n+1}(i,s) = val \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix}$ and $v_n(i,s) = val \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix}$ where $A_1 \geq A_2$, $B_1 \geq B_2$, $C_1 \geq C_2$, $D_1 \geq D_2$ from the hypothesis. Then $v_{n+1}(i,s) \geq v_n(i,s)$ and induction step goes through.

Again trivially $v_0(i+1,s) \leq v_0(i,s)$ and $v_0(i,s') \leq v_0(i,s)$ if $s' \geq s$ since all the times are zero. Assume true for $v_n(i,s)$ for all i and s then comparing

$v_{n+1}(i+1,s)$ and $v_{n+1}(i,s)$, stochastic ordering implies $\sum_{j=1}^{N} P_{i+1,j} v_n(j, \alpha s+c) \leq \sum_{j=1}^{N} P_{ij} v_n(j, \alpha s+c)$ and $\sum_{j=1}^{N} P_{i+1,j} v_n(j, \alpha s+c-1) \leq \sum_{j=1}^{N} P_{ij} v_n(j, \alpha s+c-1)$ and hence that $v_{n+1}(i+1,s) \leq v_{n+1}(i,s)$. So the induction step goes through. The same argument shows that if $v_n(i,s') \leq v_n(i,s)$ if $s \leq s'$, then $v_{n+1}(i,s') \leq v_{n+1}(i,s)$ if $s \leq s'$.

Corollary 4.3.1 $v(i,s)$ is non-increasing in i and s.

Proof Since $v_n(i,s)$ is non-deceasing in n and bounded above by T/b_1 where T is the expected number of periods to move from state 1 to state N under the do nothing strategy, then $v_n(i,s)$ is a monotonic bounded sequence and so converges to $v(i,s)$. Hence the properties, $v_n(i+1,s) \leq v_n(i,s)$, $v_n(i,s') \leq v_n(i,s)$ if $s \leq s'$ hold for the limit function $v(i,s)$.

With these results, it is possible to describe several features of the optimal strategies. Firstly we show that if the item is "down(in state N)" then player I will want to repair it, while player II will want to make the environment dangerous if they can. This ability to make the situation dangerous can only occur if $\alpha s - (1-c) \geq 0$ or $s \geq (1-c)/\alpha$. Since if player II starts with a sleep index of 0, the maximum the index can be is $s < c/(1-\alpha)$, player II can play the dangerous strategy if $c/(1-\alpha) \geq (1-c)/\alpha$, i.e. $\alpha + c \geq 1$, so if $\alpha + c < 1$, the resultant game becomes trivial with player II only able to play the peaceful strategy and the results of the 1-player situation in Chapter 3, holding.

Theorem 4.3.1 Provided $\alpha + c \geq 1$, then in state N

i) if s satisfies $s \geq (1-c)/\alpha$, the optimal strategies are "Repair vs Dangerous"

ii) if s satisfies $s < (1-c)/\alpha$, the optimal strategies are "Repair vs Peaceful"

Proof The payoff matrix in the subgame $\Gamma_{N,s}$ is

$\Gamma_{N,s}$	Making Peaceful Situation	Making Dangerous Situation
Do Nothing	$(1-b_1)(1+v(N,\alpha s+c))$	$(1-b_2)(1+v(N,\alpha s+c-1))$
Repair	$(1-b_1)(1+v(1,\alpha s+c))$	$(1-b_2)(1+v(1,\alpha s+c-1))$

Since by Corollary 4.3.1, $v(N,s) \leq v(1,s)$, it is trivial that repair strategy dominates the do nothing strategy for player I. If $s < (1-c)/\alpha$, then player II can only play the peaceful strategy and so "Repair vs Peaceful" is optimal. If $s \geq (1-c)/\alpha$, we need to show that it is better for player II to play dangerous than peaceful at the first occasion the system is in state N. Let π^* be the policy that chooses to play "peaceful" at the current situation $i = N$, and plays optimally thereafter so $v^{\pi^*_P}(N,s) = (1-b_1)(1+v(1,\alpha s+c))$. Let π_1 be the policy that plays peaceful at the current $i = N$, and the same as π^* except that at the next down situation chooses the dangerous environment(even if this policy may really be nonfeasible since the index value may subsequently go negative). Since playing dangerous rather than peaceful cannot increase the time until a catastrophic event $v^{\pi_1}(N,s) \leq v^{\pi^*_P}(N,s)$. Let π_2 be the policy that plays dangerous now and peaceful at the next down event, but otherwise chooses the strategies suggested by π_1 and π^*. Let K be the expected time between now when $i = N$ and next time $i = N$ under π^* and T is expected time until a catastrophic event from next time $i = N$ under π^* conditional on reaching the second $i = N$ event. Then

$$v^{\pi_1}(N,s) = (1-b_1)(K+(1-b_2)T) > (1-b_2)(K+(1-b_1)T) = v^{\pi_2}(N,s)$$

If π_D^* is the optimal policy for player II against the optimal policy of player I following a choice of dangerous environment now, then $v^{\pi_D^*}(N,s) \leq v^{\pi_2}(N,s) \leq v^{\pi_1}(N,s) \leq v^{\pi_P^*}(N,s)$ and it is best for player II to choose "dangerous" environment as the best response.

If the stand-by system is working then one can have any of the four combinations of pure strategies being chosen or even mixed strategies. What one can show though is that if the sleep index is so low, that player II cannot provoke a dangerous environment either this period or next period then player I will do nothing if the system is working.

Theorem 4.3.2 If $s < (1-\alpha c - c)/\alpha^2$, then player I will do nothing in state (i,s) when i is a working state($i < N$).

Proof The condition on s means that player II can only invoke a peaceful environment for the next two periods. Consider the possible strategies for player I over these next two periods,

strategy 1	:	repair in both periods
strategy 2	:	repair in period 1 and do nothing in period 2
strategy 3	:	do nothing in period 1 and repair in period 2

Let W_1, W_2, W_3 be the respective expected times until a catastrophic event if the optimal policy is used after the first two periods. Then

$$\begin{aligned}
W_1 &= (1-b_1)(1+(1-b_1)+(1-b_1)v(1,\alpha^2 s+\alpha c+c)) \\
W_2 &= (1-b_1)(1+1+\sum_{j=1}^{N} P_{ij}v(j,\alpha^2 s+\alpha c+c)) \\
W_3 &= 1+(1-b_1)+(1-b_1)v(1,\alpha^2 s+\alpha c+c)
\end{aligned}$$

and trivially $W_3 \geq W_1$ and $W_3 \geq W_2$ since $v(1,s) \geq v(j,s)$ for all j and s. Hence the do nothing now policy dominates the policies that repair now and the result holds.

It need not be the case that do nothing is optimal even if one is in the new state $i = 1$ because an opponent has to play peacefully if the sleep index is s where $\alpha s + c - 1 < 0$. Repairing keeps the item in state 1, while it could degrade under the do nothing strategy. This phenomenon is exhibited in a subsequent example($s = 0.6$ in table 4-1). Before doing that, we will show that if in the $\alpha = 1$ the system is working, and s is large enough, then either players choose do nothing vs peaceful or they play mixed strategies where player I almost always plays do nothing. To do that we need the following limit result.

Lemma 4.3.2 In the case $\alpha = 1$, as $s \to \infty$, $v_n(i,s)$ and $v(i,s)$ converge respectively to $v_n(i)$ and $v(i)$ where

$$v_n(i) = \max\{1+\sum_{j=1}^{N} P_{ij}v_{n-1}(j), (1-b_2)(1+v_{n-1}(1))\}$$

and

$$v(i) = \max\{1 + \sum_{j=1}^{N} P_{ij}v(j), (1 - b_2)(1 + v(1))\}$$

These equations correspond to the situation where player II is choosing the dangerous environment all the time.

<u>Proof</u> If T is the expected number of periods for the system to go from state 1 to state N under the transition matrix P_{ij}, then $v_n(i, s)$ must be bounded above by $\frac{T}{b_1}$ since the time until a catastrophic event under any policy must be less than a policy where the only possible times for catastrophic events are when the system is in state N in a peaceful environment(and repair is immediately undertaken). From Lemma 4.3.1 and Corollary 4.3.1, $v_n(i, s)$ and $v(i, s)$ are non-increasing sequence in s, bounded above and so must converge. In the limit since $b_1 < b_2$, player II's dangerous strategy dominates its peaceful one, since the payoffs against do nothing are the same, and against repair $(1 - b_2)(1 + v_{n-1}(1)) < (1 - b_1)(1 + v_{n-1}(1))$.

We are now in a position to describe what happens in the game when the sleep index gets very large.

<u>Theorem 4.3.3</u> In the game with $\alpha = 1$, if we are in a working state i, then for any $\varepsilon > 0$, $\exists\ S$ so that for $s \geq S$, the optimal strategies are either a) Do Nothing vs Peaceful, or b) mixed strategies where player I plays Do Nothing with a probability at least $1 - \varepsilon$.

Proof Consider the payoff matrix in the subgame $\Gamma_{i,s}^{n}$ of the game with n periods to go,

$\Gamma_{k\neq N,\ s}^{n}$	Making Peaceful Situation	Making Dangerous Situation
Do Nothing	$1+\sum_{j=k}^{N} P_{kj}v_{n-1}(j,\ s+c) : A_n$	$1+\sum_{j=k}^{N} P_{kj}v_{n-1}(j,\ s+c-1) : B_n$
Repair	$(1-b_1)(1+v_{n-1}(1,\ s+c)) : C_n$	$(1-b_2)(1+v_{n-1}(1,\ s+c-1)) : D_n$

and let A, B, C, D be the comparable values in $\Gamma_{i,s}$ when v_{n-1} is replaced by v. From Lemma 4.3.1 and the stochastic ordering property it follows that $B > A(B_n > A_n$ though as $s \to \infty$,the difference becomes very small). We now will prove $B > D$. By convergence we can choose a N and a S so that $|v_n(j,s) - v(j,s)| < \varepsilon$ for all j,s if $n \geq N$ and $|v(j,s) - v(j)| < \varepsilon$ for all j where $v_n(j)$ is defined in Lemma 4.3.2 provided $s > S$. Then,

$$
\begin{aligned}
& 1+\sum_{j=k}^{N} P_{kj}v(j,\ s) \\
\geq\ & 1+\sum_{j=k}^{N} P_{kj}v_{n+1}(j,\ s) - \varepsilon \geq 1+\sum_{j=k}^{N} P_{kj}v_{n+1}(j) - 2\varepsilon \\
\geq\ & 1+\sum_{j=k}^{N} P_{kj}(1-b_2)(1+v_n(1)) - 2\varepsilon \\
=\ & 1+(1-b_2)+(1-b_2)v_n(1) - 2\varepsilon \\
\geq\ & 1+(1-b_2)+(1-b_2)v_n(1,s) - 3\varepsilon \\
\geq\ & 1+(1-b_2)+(1-b_2)v(1,s) - 4\varepsilon \\
\geq\ & (1-b_2)+(1-b_2)v(1,s) \text{ provided } \varepsilon < \frac{1}{4}
\end{aligned}
$$

Hence $B > D$.

If $A \geq C$, then the fact $B > D$, means Do Nothing dominates Repair for player I and $A < B$ means that Peaceful dominates dangerous for player II. Thus Do Nothing vs Peaceful is optimal. If the case $A < C$, note also that if $b_2 > b_1$, for s large enough $C > D$ since

$$\begin{aligned} & (1-b_1)(1+v(1,\ s+c)) \geq (1-b_1)(1+v(1)) - \varepsilon \\ > \ & (1-b_2)(1+v(1)) + \varepsilon \geq (1-b_2)(1+v(1,s+c-1)) \end{aligned}$$

Hence with $C > A, C > D, B > A, B > D$, the optimal strategy is a mixed one with player I playing $\left(\frac{C-D}{C+B-A-D}, \frac{B-A}{C+B-A-D}\right)$. For any $\varepsilon > 0$ choose ε' so $\varepsilon' < (b_2 - b_1)\varepsilon$ and then choose S^* so for $s \geq S^*$ so that $|v(j,s) - v(j,s')| < \varepsilon'$ for $s, s' \geq S^*$ and $\varepsilon' < (b_2 - b_1)\min_{s \geq S} v(1,s)$. Then player I plays repair with probability

$$\begin{aligned} \frac{B-A}{C+B-A-D} &\leq \frac{\varepsilon'}{b_2 - b_1 - \varepsilon' + (b_2 - b_1)v(1, s+c-1)} \leq \frac{\varepsilon'}{b_2 - b_1} < \varepsilon \\ & \text{where} \\ B - A &= \sum_{j=k}^{N} P_{kj}[v(j,\ s+c-1) - v(j,\ s+c)] \\ C + B - A - D &= \sum_{j=k}^{N} P_{kj}[v(j,\ s+c-1) - v(j,\ s+c)] + b_2 - b_1 + \\ & [v(1,\ s+c) - v(1,\ s+c-1) + b_2 v(1,\ s+c-1) - b_1 v(1,\ s+c)] \end{aligned}$$

and the result holds.

4.4 Numerical Example

The actual policies in specific case can be obtained by using value iteration to do the calculations. All the following examples have three equipment states- 1(new),2(used),3(failed)- and doing nothing gives the following transition probabilities, P_{ij}

$P =$

$i \setminus j$	1	2	3
1	0.3	0.4	0.3
2	0	0.4	0.6
3	0	0	1

Assume the constraint is that $c = 0.3$ so player II can only create a dangerous environment 30% of the time. The first example is unconstrained where $\alpha = 1$.

Tables 4-1 and 4-2 give the results in the new state($i = 1$) firstly when $b_1 = 0.1$ and $b_2 = 0.5$(Table 4-1) and then when $b_1 = 0.4$ and $b_2 = 0.5$(Table 4-2). Notice that until the sleep index s is at least 0.7, player II can only choose the peaceful environment. For $s < 0.4$, player I does nothing(Theorem 4.3.2) but notice in Table 4-1 at $s = 0.6$, player I will repair, even though(perhaps because) player II can only ensure a peaceful environment. When the b_1, b_2 are quite different, the optimal strategies are mixed as s increases, though player I's probability tends to 1. When s is large enough, as Theorem 4.3.3 applies, in case 1 an $\varepsilon-$mixed strategy is optimal and in case 2 Do Nothing vs Peaceful is optimal.

Table 4-1. The Result for $i = 1(\text{new}), b_P = 0.1, b_D = 0.5$

	Entry				GV	I		II	
s	DN vs P	DN vs D	R vs P	R vs D	v	DN	R	P	D
0.0	8.545694	0	7.901159	0	8.545694	DN	–	P	–
0.1	7.991002	0	7.244038	0	7.991002	DN	–	P	–
0.2	7.846352	0	7.063115	0	7.846352	DN	–	P	–
0.3	7.779409	0	7.001468	0	7.779409	DN	–	P	–
0.4	7.049254	0	6.966484	0	7.049254	DN	–	P	–
0.5	6.848216	0	6.845989	0	6.848216	DN	–	P	–
0.6	6.771734	0	6.779698	0	6.779698	–	R	P	–
0.7	6.734876	9.258600	6.745477	4.772648	6.740826	0.44	0.56	$1-\varepsilon$	ε
0.8	6.549694	8.756620	6.663171	4.495315	6.606939	0.5	0.5	0.97	0.03
0.9	6.468176	8.610763	6.599401	4.422995	6.533275	0.5	0.5	0.97	0.03
1.0	6.429325	8.545694	6.562935	4.389533	6.495242	0.51	0.49	0.97	0.03
2.0	5.980036	6.429325	6.257458	3.646075	6.020760	0.85	0.15	0.91	0.09
3.0	5.845124	5.980036	6.152195	3.476366	5.859863	0.95	0.05	0.89	0.11
4.0	5.794211	5.845124	6.111484	3.417886	5.800096	0.98	0.02	0.88	0.12
5.0	5.774245	5.794211	6.095469	3.395269	5.776603	0.99	0.01	0.88	0.12
15.0	5.761531	5.761532	6.085306	3.380726	5.761531	$1-\varepsilon$	ε	0.88	0.12
27.0	5.761531	5.761531	6.085306	3.380726	5.761531	$1-\varepsilon$	ε	0.88	0.12
35.0	5.761531	5.761531	6.085306	3.380726	5.761531	$1-\varepsilon$	ε	0.88	0.12

Table 4-2. The Result for $i = 1$(new), $b_P = 0.4, b_D = 0.5$

	Entry				GV	I		II	
s	DN vs P	DN vs D	R vs P	R vs D	v	DN	R	P	D
0.0	6.231987	0	4.212781	0	6.231987	DN	–	P	–
0.1	6.079537	0	4.128808	0	6.079537	DN	–	P	–
0.2	6.036238	0	4.100696	0	6.036238	DN	–	P	–
0.3	6.021390	0	4.090763	0	6.021390	DN	–	P	–
0.4	5.881433	0	4.087023	0	5.881433	DN	–	P	–
0.5	5.834579	0	4.073457	0	5.834579	DN	–	P	–
0.6	5.818203	0	4.066845	0	5.818203	DN	–	P	–
0.7	5.811789	6.394473	4.064020	3.615947	5.811789	DN	–	P	–
0.8	5.789178	6.278914	4.062832	3.539723	5.789178	DN	–	P	–
0.9	5.778158	6.244166	4.060517	3.518074	5.778158	DN	–	P	–
1.0	5.773448	6.231987	4.059103	3.510651	5.773448	DN	–	P	–
2.0	5.761807	5.773448	4.056922	3.382586	5.761807	DN	–	P	–
3.0	5.761536	5.761807	4.056872	3.380768	5.761807	DN	–	P	–
4.0	5.761531	5.761535	4.056871	3.380727	5.761531	DN	–	P	–
5.0	5.761531	5.761531	4.056871	3.380726	5.761531	DN	–	P	–
15.0	5.761531	5.761531	4.056871	3.380726	5.761531	DN	–	P	–
27.0	5.761531	5.761531	4.056871	3.380726	5.761531	DN	–	P	–
35.0	5.761531	5.761531	4.056871	3.380726	5.761531	DN	–	P	–

Tables 4-3 and 4-4 are the policies for the used and failed states in the case when $b_1 = 0.1$ and $b_2 = 0.5$. In state 2, one has Do Nothing vs Peaceful for $s < 0.4$(no dangerous environments for at least two periods), then one has Repair vs Peaceful, at $0.4 \leq s < 0.7$. The mixed strategies are optimal as s increases and as $s \longrightarrow \infty$, player I tends to do nothing with probability $1 - \varepsilon$ while player II tends to $(0.62, 0.38)$. Table 4-4 confirms the results of Theorem 4.3.1 that when the unit is down it must be repaired and the enemy will seek to make the environment dangerous if he can. Looking at the same problem $b_1 = 0.1$, $b_2 = 0.5$ but in the discounted case with $\alpha = 0.8$ and $c = 0.4$(not 0.3) leads to Tables 4-5 and 4-6.

Table 4-3. The Result for $i = 2$(not new, but working), $b_P = 0.1, b_D = 0.5$

	Entry				GV	I		II	
s	DN vs P	DN vs D	R vs P	R vs D	v	DN	R	P	D
0.0	8.312322	–	7.901159	–	8.312322	DN	–	P	–
0.1	7.966175	–	7.244038	–	7.966175	DN	–	P	–
0.2	7.845686	–	7.063115	–	7.845686	DN	–	P	–
0.3	7.779409	–	7.001468	–	7.779409	DN	–	P	–
0.4	6.458816	–	6.966484	–	6.966484	–	R	P	–
0.5	60214748	–	6.845989	–	6.845989	–	R	P	–
0.6	6.138669	–	6.779698	–	6.779698	–	R	P	–
0.7	6.103182	7.065249	6.745477	4.772648	6.488706	0.40	0.60	0.87	0.13
0.8	5.835919	8.532539	6.663171	4.495315	6.294506	0.45	0.55	0.83	0.17
0.9	5.745506	8.375801	6.599401	4.42995	6.212769	0.45	0.55	0.82	0.18
1.0	5.708	8.312322	6.562935	4.389533	6.174224	0.45	0.55	0.82	0.18
2.0	5.265155	5.708540	6.257458	3.646075	5.409183	0.85	0.15	0.68	0.32
3.0	5.130156	5.265155	6.152195	3.476366	5.179242	0.95	0.05	0.64	0.36
4.0	5.079677	5.130156	6.111484	3.417886	5.098657	0.98	0.02	0.62	0.38
5.0	5.059901	5.079677	6.095469	3.395269	5.067430	0.99	0.01	0.62	0.38
15.0	5.047299	5.047299	6.085306	3.380726	5.047299	$1-\varepsilon$	ε	0.62	0.38
27.0	5.047299	5.047299	6.085306	3.380726	5.047299	$1-\varepsilon$	ε	0.62	0.38
35.0	5.047299	5.047299	6.085306	3.380726	5.047299	$1-\varepsilon$	ε	0.62	0.38

Table 4-4. The Result for $i = 3(\text{down}), b_P = 0.1, b_D = 0.5$

s	Entry				GV	I		II	
	DN *vs* P	DN *vs* D	R *vs* P	R *vs* D	v	DN	R	P	D
0.0	7.201043	–	7.901159	–	7.901159	–	R	P	–
0.1	7.169558	–	7.244038	–	7.244038	–	R	P	–
0.2	7.061118	–	7.063115	–	7.063115	–	R	P	–
0.3	7.001468	–	7.001468	–	7.001468	–	R	P	–
0.4	5.195171	–	6.966484	–	6.966484	–	R	P	–
0.5	4.945585	–	6.845989	–	6.845989	–	R	P	–
0.6	4.880504	–	6.779698	–	6.779698	–	R	P	–
0.7	4.850395	4.450395	6.745477	4.772648	4.772648	–	R	–	D
0.8	4.521847	4.121847	6.663171	4.495315	4.495315	–	R	–	D
0.9	4.431391	4.031391	6.599401	4.442995	4.422995	–	R	–	D
1.0	4.400580	4.000580	6.562935	4.389533	4.389533	–	R	–	D
2.0	4.112814	2.444766	6.257458	3.646075	3.646075	–	R	–	D
3.0	4.007289	2.284897	6.152195	3.476366	3.476366	–	R	–	D
4.0	3.967883	2.226272	6.111484	3.417886	3.417886	–	R	–	D
5.0	3.952436	2.204379	6.095469	3.395269	3.395269	–	R	–	D
15.0	3.942610	2.190339	6.085306	3.380726	3.380726	–	R	–	D
27.0	3.942610	2.190339	6.085306	3.380726	3.380726	–	R	–	D
35.0	3.942610	2.190339	6.085306	3.380726	3.380726	–	R	–	D

Table 4-5. The Result for $i = 1(\text{new}), c = 0.4, \alpha = 0.8$

s	Entry				GameValue	I		II	
	DN vs P	DN vs D	R vs P	R vs D	v	DN	R	P	D
0.0	9.119139	0	8.433221	0	9.119139	DN	–	P	–
0.1	8.499747	0	7.655538	0	8.499747	DN	–	P	–
0.2	8.480989	0	7.632890	0	8.480989	DN	–	P	–
0.3	8.436260	0	7.592634	0	8.436260	DN	–	P	–
0.4	8.390626	0	7.551563	0	8.390626	DN	–	P	–
0.5	7.517332	0	7.513618	0	7.517332	DN	–	P	–
0.6	7.390844	0	7.492803	0	7.492803	–	R	P	–
0.7	7.332755	0	7.409354	0	7.409354	–	R	P	–
0.8	7.327393	9.760024	7.403906	5.030051	7.366117	0.49	0.51	0.98	0.02
0.9	7.296601	9.222378	7.371813	4.738161	7.328369	0.58	0.42	0.98	0.02
1.0	7.214618	9.206344	7.289975	4.730074	7.247593	0.56	0.44	0.98	0.02
1.1	7.196503	9.206344	7.263204	4.719848	7.225783	0.56	0.44	0.99	0.01
1.2	7.021900	9.120808	7.231623	4.685983	7.116675	0.55	0.45	0.95	0.05
1.3	6.990664	8.940034	7.190841	4.598464	7.076582	0.57	0.43	0.96	0.04
1.4	6.966562	8.490526	7.159423	4.245263	7.032787	0.66	0.34	0.96	0.04
1.5	6.928468	8.474689	7.122899	4.237344	6.996304	0.65	0.35	0.96	0.04
1.6	6.872290	8.392972	7.038725	4.196486	6.930300	0.65	0.35	0.96	0.04
1.7	6.743334	7.525998	6.983073	4.175824	6.7956.1	0.78	0.22	0.93	0.07
1.8	6.713779	7.396593	6.963334	4.167040	6.762757	0.80	0.20	0.93	0.07
1.9	6.682847	7.351381	6.944154	4.130449	6.733013	0.81	0.19	0.92	0.08
2.0	6.680185	7.330340	6.942264	4.115207	6.729240	0.81	0.19	0.92	0.08

Table 4-6. The Result for $i = 3$(down), $c = 0.4, \alpha = 0.8$

s	Entry				GameValue	I		II	
	DN vs P	DN vs D	R vs P	R vs D	v	DN	R	P	D
0.0	7.679899	0	8.433221	0	8.433221	–	R	P	–
0.1	7.647301	0	7.655538	0	7.655538	–	R	P	–
0.2	7.632890	0	7.632890	0	7.632890	–	R	P	–
0.3	7.592634	0	7.592634	0	7.592634	–	R	P	–
0.4	7.551563	0	7.551563	0	7.551563	–	R	P	–
0.5	5.415381	0	7.513618	0	7.513618	–	R	P	–
0.6	5.154424	0	7.492803	0	7.492803	–	R	P	–
0.7	5.148572	0	7.409353	0	7.409353	–	R	P	–
0.8	5.144705	4.690239	7.403906	5.030051	5.030051	–	R	–	D
0.9	5.128855	4.315922	7.371813	4.738161	4.738161	–	R	–	D
1.0	5.106849	4.307067	7.289975	4.730074	4.730074	–	R	–	D
1.1	5.105788	4.297863	7.263204	4.719848	4.719848	–	R	–	D
1.2	4.717435	4.267385	7.231623	4.685983	4.685983	–	R	–	D
1.3	4.707608	4.262756	7.190841	4.598464	4.598464	–	R	–	D
1.4	4.700986	4.245263	7.159423	4.245263	4.245263	–	R	–	D
1.5	4.667655	4.237344	7.122899	4.237344	4.237344	–	R	–	D
1.6	6.872290	4.196486	7.038725	4.196486	4.196486	–	R	–	D
1.7	6.743334	3.017556	6.983073	4.175824	4.175824	–	R	–	D
1.8	6.713779	2.863917	6.963334	4.167040	4.167040	–	R	–	D
1.9	6.682847	2.862228	6.944154	4.130449	4.130449	–	R	–	D
2.0	6.680185	2.859057	6.942564	4.115207	4.115207	–	R	–	D

<table>
<tr><th>$i\backslash s$</th><th>0.0</th><th>0.1</th><th>0.2</th><th>0.3</th><th>0.4</th><th>0.5</th><th>0.6</th><th>0.7</th><th>0.8</th><th>0.9</th><th>1.0</th><th>5.0</th><th>35.0</th></tr>
<tr><td>1</td><td colspan="6">DN v P</td><td>RvP</td><td colspan="6">Mixed</td></tr>
<tr><td>2</td><td colspan="4">DN v P</td><td colspan="3">RvP</td><td colspan="6">Mixed</td></tr>
<tr><td>3</td><td colspan="7">R v P</td><td colspan="6">R v D</td></tr>
</table>

Figure 4-1. Simple form of Table 4-1, 4-3, 4-4.

<table>
<tr><th>$i\backslash s$</th><th>0.0</th><th>0.1</th><th>0.2</th><th>0.3</th><th>0.4</th><th>0.5</th><th>0.6</th><th>0.7</th><th>0.8</th><th>0.9</th><th>1.0</th><th>5.0</th><th>35.0</th></tr>
<tr><td>1</td><td colspan="6">DN v P</td><td colspan="2">RvP</td><td colspan="5">Mixed</td></tr>
<tr><td>3</td><td colspan="8">R v P</td><td colspan="5">R v D</td></tr>
</table>

Figure 4-2. Simple form of Table 4-5, 4-6.

Again Table 4-6 confirms the results of Theorem 4.3.1, since player II can only play dangerous if $s \geq 0.75$, while Table 4-5 shows the changes as the sleep index increase from Do Nothing vs Peaceful to Repair vs Peaceful to mixed strategies. Note that 2 is the greatest value the sleep index can be ($\frac{c}{1-\alpha}$ when $c = 0.4$ and $\alpha = 0.8$), and in this case both players are playing genuinely mixed strategies.

4.5 Model with Local Constraints on Effort

4.5.1 Introduction In the two models of section three a rest period by the enemy continued to have a positive effect on their performance as far into the future as one wanted to consider, even if in the discounted case, the effect diminished over time. An alternative is to say that there is a definite time horizon T on the effect a peaceful period(sleep period) can have. Thus the obvious local analogy to the constraints in section three is to say that out of T consecutive periods, player II can only choose a dangerous environment in at most C of them.

4.5.2 Game Model In the game that models this problem, the subgames $\Gamma_{i,X}$ are indexed by the current state of the equipment i, and the description of the environment in the last T periods X, and the value of the game starting in such a subgame is $v(i,x)$. The state space for X consists of the 2^T possible sequences of P and D(peaceful and Dangerous), $X = X_1...X_T$, and in $\Gamma_{i,X}$ if player II choose Peaceful, the game moves to the state $X' = X_2...X_TP$ while if he chooses dangerous, the game moves to state $X'' = X_2...X_TD$, so the payoff matrix in the subgame $\Gamma_{i,X}$ is

$\Gamma_{i,X}$	Making Peaceful Situation	Making Dangerous Situation
DN	$(1-\delta_{iN}b_1)(1+\sum_{j=1}^{N} P_{ij}v(j,X_2...X_TP))$	$(1-\delta_{iN}b_2)(1+\sum_{j=1}^{N} P_{ij}v(j,X_2...X_TD))$
R	$(1-b_1)(1+v(1,X_2...X_TP))$	$(1-b_2)(1+v(1,X_2...X_TD))$

(4.7)

where $\delta_{iN} = 1$ if $i = N$, 0 otherwise.

The peaceful strategy for player II is always allowed but the dangerous strategy is only possible if the number of D's in $X_2...X_TD$ is less than or equal to C. For example if $T = 2$ and $C = 2$ there are four states PP, PD, DP and DD and one can choose peaceful or dangerous in any of them. However if $1 \leq C < 2$, then there are only three states allowed PP, PD and DP and one can only choose dangerous in PP or DP. In the case $T = 3$ if $1 \leq C < 2$, then there are 4 states PPP, DPP, PDP and PPD and the dangerous strategy can be played in the first two states, but not in the second two. For the general T and C, where $C = \lfloor C \rfloor$, there are $\sum_{C=0}^{C} {}^{T}C_{C}$ environmental states and in ${}^{T-1}C_{C}$ of these one cannot choose the dangerous option(where there are C dangerous states in the last $T-1$ periods). As C increases the game approaches the unconstrained game as the following lemma formally indicated.

Lemma 4.5.1 If $C = T$, then $v(i, X) = v(i)$(defined by (4.1) and (4.2)) and the optimal strategy is given by Theorem 4.2.1.

Proof Since $C = T$, there are no restriction on when player II can play dangerous and the payoff matrix (4.7) reduces to that of (4.1), (4.2).

Another obvious observation is that the values $v(i, X_1X_2...X_T)$ are independent of X_1, since it drops out of the history at the next period. There are other fairly obvious relationships between the values $v(i, X)$ for different X.

Theorem 4.5.1 Define a partial ordering $\succ$ on X by

a) $Y \succ X$ if for some k $X_k = P, Y_k = D, X_l = Y_l, l \neq k$

b) $Y \succ X$ if for some $l > k$ $X_l = Y_k = P, X_k = Y_l = D, X_m = Y_m, m \neq k, l$,

then if $Y \succ X$, $v(i, Y) \geq v(i, X)$.

Proof Consider the games $v_n(i, X)$ which last only for n periods and which converge to $v(i, X)$ as $n \to \infty$ by same approach as Lemma 4.2.1. Trivially if

$X \prec Y, v_0(i,X) = 0 \leq v_0(i,Y) = 0$, so assume the inequalities hold for $v_n(i,X)$. For $\Gamma^{n+1}_{i,X}, X_k = P, Y_k = D, X_l = Y_l$, then $X' = (X_2...X_TP) \prec (Y_2...Y_TP) = Y'$, $X'' = (X_2...X_TD) \prec (Y_2...Y_TD) = Y''$ or if $k = 1$, $X' = Y'$, $X'' = Y''$. In all case $v_n(i,X') \leq v_n(i,Y')$, $v_n(i,X'') \leq v_n(i,Y'')$ so each entry on the payoff matrix for $v_{n+1}(i,X)$ in equation (4.7) is smaller than the corresponding entry for $v_{n+1}(i,Y)$ and hence $v_n(i,X) \leq v_n(i,Y)$. Thus induction step is proved and result holds.

A similar proof holds if $X \prec Y$ using b) because the X', Y' at the next stage will either have $X' \prec Y'$ because of b) if $k > 1$ or a) if $k = 1$.

Thus Theorem 4.4.1 and the observation that the first time period value does not affect the value, shows that if $T = 2$ and $C = 1$, then $v(i,PP) = v(i,DP) \leq v(i,PD)$. In the case when $T = 3$ and $C = 2$, one has

$$\begin{aligned} & v(i,PPP) \\ = \; & v(i,DPP) = v(i,PDP) = v(i,DDP) \\ \leq \; & v(i,DPD) = v(i,PPD) \leq v(i,PDD) \end{aligned}$$

As the state of the stand-by system worsens, the time until a catastrophic event decreases as the following theorem in depth.

Theorem 4.5.2 For any feasible X, $v(i,X)$ is a non-increasing function in i.

Proof The proof is analogues to that in Lemma 4.2.1 using induction on n in the games $\Gamma^n_{i,X}$.

If the standby unit is down($i = N$), then player I will always wish to repair it while player II will want to make the situation dangerous if he has the resources to do so.

Theorem 4.5.3 For any C, T and state (N, X) player I will play repair while player II will play dangerous if $(X_2...X_nD)$ is feasible.

Proof In state N, the do nothing action in equation (4.7) for $v(i, X_1X_2...X_T)$ gives values

$$(1 - b_1)(1 + v(N, X_2...X_TP)) \text{ and } (1 - b_2)(1 + v(N, X_2...X_TD))$$

while the repair action gives values

$$(1 - b_1)(1 + v(1, X_2...X_TP)) \text{ and } (1 - b_2)(1 + v(1, X_2...X_TD))$$

, so Theorem 4.5.2 implies repair dominates do nothing. To show that player II will want to play dangerous rather than peaceful if that is possible we need to prove

$$(1 - b_2)(1 + v(1, X_2...X_TD)) \leq (1 - b_1)(1 + v(1, X_2...X_TP))$$

Let π^* be the optimal policy for player II from state $(1, X_2...X_TP)$ i.e. (N, X) playing P at the next period and let π_1 and π_2 identical policies except that π_1 plays P at the next time $i = N$, while π_2 plays D at this occasion. So $v^{\pi^*}(N, X) = \min\{v^{\pi_1}(N, X), v^{\pi_2}(N, X)\}$. Compare these policies with a third policy which is the same as π_1 and π_2 except it plays Dangerous at this current period when it is in (N, X) and will play P at the next time $i = N$. Then

$$\begin{aligned} & v^{\pi_3}(N, X) \\ = \ & (1-b_2)(T+(1-b_1)v(1, Y)) \\ \leq \ & (1-b_1)(T+(1-b_2)v(1, Y)) \\ = \ & v^{\pi_2}(N, X) \end{aligned}$$

Similarly

$$\begin{aligned} & v^{\pi_3}(N, X) \\ = \ & (1-b_2)(T+(1-b_1)v(1, Y)) \\ \leq \ & (1-b_1)(T+(1-b_1)v(1, Y)) \\ = \ & v^{\pi_1}(N, X) \end{aligned}$$

Hence if $\pi^*(D)$ is optimal policy for player II starting in (N, X) and paying D first we have

$$\begin{aligned} & (1-b_2)(1+v(1, X_2...X_T D)) \\ = \ & v^{\pi^*(D)}(N, X) \leq v^{\pi_3}(N, X) \leq \min\{v^{\pi_1}(N, X), v^{\pi_2}(N, X)\} \\ = \ & v^{\pi^*}(N, X) = (1-b_1)(1+v(1, X_2...X_T P)) \end{aligned}$$

and the result holds.

Finally we can show that for a working state (i, X) there is a limit state $i^*(X)$ so that if $i < i^*(X)$ one does nothing and otherwise one repairs.

Theorem 4.5.4 For any C, T and state(i, X) $i < N$ there exists $i_1^*(X), i_2^*(X)$ so that for $i < i_1^*(X)$ player I does nothing and $i \geq i_2^*(X)$, player I does maintenance/repair whereas from $i_1^*(X)$ to $i_2^*(X) - 1$ it plays a mixed strategy.

Proof From equation (4.7) one can see that payoff of the maintenance/repair action is independent of i against peaceful and repair. Theorem 4.5.2 and stochastic ordering property means that the do nothing action in state (i, X) are non-increasing in i, both against peaceful and dangerous. Hence for any X define

$$i^*(P, X) = \min\{i|(1 - b_1)(1 + v(1, X_2...X_T P)) > 1 + \sum_{j=1}^{N} P_{ij} v(j, X_2...X_T P)\}$$

and

$$i^*(D, X) = \min\{i|(1 - b_2)(1 + v(1, X_2...X_T D)) > 1 + \sum_{j=1}^{N} P_{ij} v(j, X_2...X_T D)\}$$

Then if $i < i_1^*(X) = \min\{i^*(P, X), i^*(D, X)\}$, the do nothing action dominates the repair action in both peaceful and dangerous environment, and hence is optimal.

If $i \geq i_2^*(X) = \max\{i^*(P, X), i^*(D, X)\}$, the repair/maintenance action dominates the do nothing action in both peaceful and dangerous situations and so is optimal.

If $i_1^*(X) \leq i < i_2^*(X)$, then there is no domination between player I's actions and so a mixed strategy is optimal.

4.6 Numerical Example

To show these results in action consider the following numerical example where $T = 2, C = 1$, so the possible states are PP, DP and PD, where player II cannot play dangerous in the last state. The transition matrices of the state of the equipment when nothing is done is

$P_{ij} =$

$i\backslash j$	1	2	3	4	5	6	7	8	9	10
1	0.2	0.2	0.2	0.1	0.08	0.05	0.05	0.05	0.05	0.02
2	0	0.2	0.2	0.2	0.1	0.1	0.08	0.05	0.04	0.03
3	0	0	0.2	0.2	0.2	0.1	0.1	0.1	0.05	0.05
4	0	0	0	0.2	0.2	0.2	0.15	0.1	0.1	0.05
5	0	0	0	0	0.2	0.3	0.2	0.1	0.1	0.1
6	0	0	0	0	0	0.2	0.3	0.2	0.2	0.1
7	0	0	0	0	0	0	0.2	0.3	0.3	0.2
8	0	0	0	0	0	0	0	0.3	0.4	0.3
9	0	0	0	0	0	0	0	0	0.4	0.6
10	0	0	0	0	0	0	0	0	0	1

In our first example $b_1 = 0.1, b_2 = 0.5$, recall that $v(\cdot, PP) = v(\cdot, DP) \leq v(\cdot, PD)$.

Table 4-7. PP and DP cases, $b_1 = 0.1, b_2 = 0.5$

i	Entry				GV	I		II	
	DN vs P	DN vs D	R vs P	R vs D	v	DN	R	P	D
1	15.6105	16.0975	14.9129	8.2849	15.6105	DN	–	P	–
2	15.3212	15.9583	14.9129	8.2849	15.3212	DN	–	P	–
3	14.9274	15.8768	14.9129	8.2849	14.9274	DN	–	P	–
4	14.6736	15.8738	14.9129	8.2849	14.7103	0.85	0.15	0.97	0.03
5	14.2448	15.8738	14.9129	8.2849	14.3766	0.80	0.20	0.92	0.08
6	13.9207	15.8738	14.9129	8.2849	14.1465	0.77	0.23	0.88	0.12
7	13.1537	15.8738	14.9129	8.2849	13.6656	0.71	0.29	0.81	0.19
8	12.5169	15.8738	14.9129	8.2849	13.3224	0.66	0.34	0.76	0.24
9	11.0095	15.8738	14.9129	8.2849	12.6617	0.58	0.42	0.66	0.34
10	8.3369	7.9369	14.9129	8.2849	8.2849	–	R	–	D

Table 4-8. PD case, $b_1 = 0.1, b_2 = 0.5$

	Entry				GV	I		II	
i	$DN\ vs\ P$	$DN\ vs\ D$	$R\ vs\ P$	$R\ vs\ D$	v	DN	R	P	D
1	15.6105	–	14.9129	–	15.6105	DN	–	P	–
2	15.3212	–	14.9129	–	15.3212	DN	–	P	–
3	14.9274	–	14.9129	–	14.9274	DN	–	P	–
4	14.6736	–	14.9129	–	14.9129	–	R	P	–
5	14.2448	–	14.9129	–	14.9129	–	R	P	–
6	13.9207	–	14.9129	–	14.9129	–	R	P	–
7	13.1537	–	14.9129	–	14.9129	–	R	P	–
8	12.5169	–	14.9129	–	14.9129	–	R	P	–
9	11.0095	–	14.9129	–	14.9129	–	R	P	–
10	8.3369	–	14.9129	–	14.9129	–	R	P	–

$X\backslash i$	1	2	3	4	5	6	7	8	9	10
PP, DP	DNvP			Mix						RvD
PD	DNvP			RvP						

Figure 4-3. Simple form of Table 4-7, 4-8

In the PP state, if the equipment is in a good state, one plays do nothing against peaceful, while in less good states one plays mixed strategies with increasing emphasis on repair and dangerous until in the down state, one jumps to repair vs dangerous. In state PD, the expected time until catastrophic event is higher($v(i, PP) < v(i, PD)$) and $i_1^* = i_2^*$ in that as the state worsens one goes directly from do nothing vs peaceful to repair vs peaceful(dangerous not being allowed). If $b_1 = 0.1$ and $b_2 = 0.9$, so there is very high possibility of an "attack(initiating event)" in the dangerous state, one starts the mixed actions earlier as Table 4-9 shows.

Table 4-9. PP and DP cases, $b_1 = 0.1, b_2 = 0.9$

	Entry				GV	I		II	
i	DN vs P	DN vs D	R vs P	R vs D	v	DN	R	P	D
1	9.8655	10.7890	9.7751	1.0861	9.8655	DN	–	P	–
2	9.5668	10.7709	9.7751	1.0861	9.5922	0.88	0.12	0.98	0.02
3	9.1302	10.7709	9.7751	1.0861	9.2326	0.84	0.16	0.94	0.06
4	8.8479	10.7709	9.7751	1.0861	9.0159	0.82	0.18	0.91	0.09
5	8.3314	10.7709	9.7751	1.0861	8.6479	0.78	0.22	0.87	0.13
6	7.9540	10.7709	9.7751	1.0861	8.3999	0.76	0.24	0.84	0.16
7	7.0064	10.7709	9.7751	1.0861	7.8433	0.70	0.30	0.78	0.22
8	6.2106	10.7709	9.7751	1.0861	7.4375	0.66	0.34	0.73	0.27
9	4.3059	10.7709	9.7751	1.0861	6.6392	0.57	0.43	0.64	0.36
10	1.8770	1.0770	9.7751	1.0861	1.0861	–	R	–	D

Table 4-10. PD case, $b_1 = 0.1, b_2 = 0.9$

i	Entry				GV	I		II	
	DN vs P	DN vs D	R vs P	R vs D	v	DN	R	P	D
1	9.8655	–	9.7751	–	9.8655	DN	–	P	–
2	9.5668	–	9.7751	–	9.7751	–	R	P	–
3	9.1302	–	9.7751	–	9.7751	–	R	P	–
4	8.8479	–	9.7751	–	9.7751	–	R	P	–
5	8.3314	–	9.7751	–	9.7751	–	R	P	–
6	7.9540	–	9.7751	–	9.7751	–	R	P	–
7	7.0064	–	9.7751	–	9.7751	–	R	P	–
8	6.2106	–	9.7751	–	9.7751	–	R	P	–
9	4.3059	–	9.7751	–	9.7751	–	R	P	–
10	1.8770	–	9.7751	–	9.7751	–	R	P	–

$X\backslash i$	1	2	3	4	5	6	7	8	9	10
PP, DP	DNvP	Mix								RvD
PD	DNvP	RvP								

Figure 4-4. Simple form of Table 4-9, 4-10

4.7 Conclusion

These models have sought to investigate the maintenance and repair policy for a stand-by system where the environment of when it is needed is controlled by an opponent. The most obvious context for this problem is the military one either in conventional or peace keeping roles. It would also apply to emergency services needed to respond to terrorist threats, but it could also be useful for more routinely used equipment in that context such as airport passenger and luggage screening devices. We have shown that if there is no limit on resources available to the "enemy", then the problem reduces to a single decision making in a constant high risk environment. If more realistically the enemy cannot always be ready to act, but needs time to recuperate, resupply and plan, the situation is much more complex, both in the situation where the restful periods have a long term effect or only a short term effect.

One interesting feature is that the optimal policies may well be mixed so each period there is a certain probability one should perform maintenance, and a certain probability one does nothing. Clearly if there are a number of such stand-by units, the mixed policy can translate into what proportion should be given preventive maintenance at that time. If the difference between the benign and the dangerous environment (b_1, b_2) is small, one tends only to perform maintenance when system is close to failure, but in other situations one will maintain the system in a good state because one feels the environment is likely soon to be dangerous(especially if the sleep index is high). One always repairs a failed unit, no matter what the environment, but you can be sure that the enemy will seek to take advantage of the failure by increasing the danger in the environment in this case. The models introduced in this paper are the first to address the question of maintenance in an environment where failure can be catastrophic and where there is an enemy seeking such catastrophes. Clearly more sophisticated models can be developed but we believe this paper has indicated that one can get useful insights by addressing the problem as a stochastic game.

CHAPTER 5

Repair Strategies in an Uncertain Environment: Markov Decision Process Approach considering Training Factor

5.1 Introduction

In last two chapters, we have developed repair strategies of the stand-by unit which maximise the time until a catastrophic event. In the chapter 3, we examined the case where the need for the equipment varies overtime according to a Markov chain. This means that the environment can be in different states, each with their own probability of the initiating event occurring. We described the form of the optimal policy under this uncertain environment by Markov Decision Process. In the chapter 4, we look at conflict situations where the environment is controlled by an opponent. In this case the opponent's actions force the need for the equipment, and this situation is modelled as a stochastic game.

In this chapter our research is also concerned with developing repair and training strategies which maximise the time until a catastrophic event for standby units in an uncertain environment. This is extension of previous Markov decision process model. Equipment can only be used if it is in an operable state and if its users have had sufficient recent training with it. Thus as well as repairing and maintaining the equipment, it is necessary to train users. This is particularly clear in the military context where soldiers are constantly trained to operate the equipment satisfactorily under all conditions. However, a problem with training is that it increases the wear and tear of the stand-by unit even though it enhances the operator's ability to respond well to an initiating event. Another problem in the military context is that the training may be done away from where the equipment may be needed and so there is not time to move it between the training area and the front line say. In this chapter we look at the interaction between the need for training and the need to service the equipment. We develop discrete time Markov decision process formulations of the problem in order to investigate the form of the optimal action policies which maximise the expected survival time until a catastrophic event. Apart from some general discussion that training improves the skill level of the operators and so could reduce failures, none of the previous research in maintenance addresses the issue of how does the training of the operators affect the readiness of the unit. This chapter considers this issue.

We develop a Markov decision process model with random loss of learning in the training level in section 5.2. Numerical examples of these results are presented in section 5.3. In section 5.4 we examine a modified model where the effect of training does wear off, and look at when one should train, and when one should repair as a function of the environmental situation, the training level and the state of the equipment.

5.2 Training Model with Random Loss of Expertise

5.2.1 Introduction In this model, there are several environmental situations which are graded from very dangerous to completely peaceful. Each environmental situation has its own probability of an initiating event occurring which increases as the situation gets more dangerous. There are three actions for this model which means that the operator chooses among doing nothing, repairing, and training. At the end of this section we look at the special case in which we only consider do nothing and training. This corresponds to equipment which cannot be repaired though we do not consider the problem of when to replace such equipment.

5.2.2 Terminology

Possible Standby Unit Quality State, i Regular inspection of the standby unit gives information on the operation quality state of the units. We assume the standby unit has N different unit quality states, i.e. $1, 2, \cdots, N$ where state 1 means that the standby unit is like new. The state $N-1$ means that it is in a poor but still operable state, while in state N, it is in a "down" condition which means that it will not work.

The Quality State Transition Probability Matrix(QSTPM), P_{ij} When the standby unit is in quality state i at the current stage, there is a probability, P_{ij} that it will be in state j at the next period where $i, j = 1, 2, \cdots, N$ and

$$\sum_{j=1}^{N} P_{ij} = 1, \text{ where } i = 1, 2, \cdots, N-1, N$$

We assume that the QSTPM satisfies a first order stochastic ordering condition so that $\sum_{j<k} P_{ij} \geq \sum_{j<k} P_{(i+1)j}$. We assume $P_{NN} = 1$ so once the standby unit reaches the "down" state N, it remains "down" until either it is repaired, or a catastrophic event occurs. There are several situations where equipment is classified as new, excellent condition, operable, failed and regular inspection of the equipment allows one to collect data to estimate the transition probabilities, P_{ij}.

<u>Possible Environmental Situation, m</u> We assume that there are M different environmental states, $1, 2, \cdots, M-1, M$. Environmental state 1 reflects the most peaceful environment in which there is the smallest probability, b_1 of an initiating event occurring. On the other hand, environmental state M is the most dangerous state with the highest probability, b_M of an initiating event occurring. We assume b_m is non-decreasing in the index of the environmental state m and $0 \leq b_m \leq 1$. These correspond to military states of readiness, such as the US DEFCON, or the UK, black/red/amber.

<u>Environment Situation Transition Probability Matrix(ESTPM), $S_{mm'}$</u> The dynamics of the environmental situation is also described by a Markov chain with Environment Situation Transition Probability Matrix(ESTPM), $S_{mm'}$. If the environmental situation is m, $1 \leq m \leq M$ in the current stage, this changes to another environmental situation $m', 1 \leq m' \leq M$ with probability $S_{mm'}$ at the next stage, where

$$\sum_{m'=1}^{M} S_{mm'} = 1, \text{ with } m \text{ and } m' = 1, 2, \cdots, M-1, M$$

We assume the ESTPM also satisfies a first order stochastic ordering property so $\sum_{m'=1}^{k} S_{mm'} \geq \sum_{m'=1}^{k} S_{(m+1)m'}$ for any $m = 1, 2, \cdots, M-1$. The data for this can be obtained by historical analysis.

The Possible Actions There are three possible actions at each period, (1) do nothing, (2) repair and (3) training. The "do nothing" action means neither repair/maintenance nor training is undertaken. It is assumed the "repair" action which can be maintenance action, if the unit is still operable, but is a true repair in state N takes 1 unit time period. This action is not perfect in that there is a probability R_r the unit will be in quality state r after the "repair" where $\sum_{r=1}^{N} R_r = 1$. If an initiating event occurs during repair period, the standby unit cannot respond to it, and so a catastrophic event occurs automatically.

Training Level, k The operator of the stand-by unit has L different training levels, i.e. $1, 2, \cdots, L$ where training level 1 is the best training level and training level L is the worst training level. If there is no training at the moment, the training level k goes to $k^{'}$ at the next time stage with probability of $T_{kk'}$ which is called training level transition probability matrix(TLTPM). We assume there is no spontaneous improvement in training i.e. $T_{kk'} = 0$ if $k^{'} < k$, and our model would allow the deterministic decrease in operator performance $T_{k,k+1} = 1$. Training may not be ideal and could be counter productive in that after a training exercise(which we also assume to take one period) the quality of training is p with probability w_p where $p = 1, 2, \cdots, L-1, L$. The repair times and hence unit times in the model are often quite small and in this case $T_{k,k}$ would be close to 1. Training causes wear and tear on the equipment to a different extent than when it is not being used. Hence there is a transition probability matrix for the standby unit quality state variation caused by the training which is called wear and tear transition probability matrix(WTTPM), $\widetilde{P}_{ij}$. If $j < i$, $\widetilde{P}_{ij} = 0$. We assume that the WTTPM also satisfies a first order stochastic ordering condition so that $\sum_{j<k} \widetilde{P}_{ij} \geq \sum_{j<k} \widetilde{P}_{(i+1)j}$. We assume the wear and tear caused by training is more than the wear and tear caused by natural conditions and so require $\sum_{j \leq l} P_{ij} \geq \sum_{j \leq l} \widetilde{P}_{ij}$ where l is arbitrary quality state. If the unit is being repaired, no training is possible.

Catastrophic Event If an initiating event occurs either when the standby unit is down(in state N) or being repaired, a catastrophic event comes. To allow for the possibility that training could be aborted when an initiating event occurs, but only if the equipment is close to where it is needed, we say that if training is occurring there is a probability $(1 - t)$, the equipment can respond to the initiating event. Finally at an initiating event it is not enough for the equipment to be operating, but the training must be of a sufficient quality if there is not to be a catastrophic outcome. We assume that with training level k, one cannot successfully respond to an initiating event with probability K_k where K_k increases with k and $K_1 = 0$ and $K_L = 1$.

5.2.3 Model

Model State Space The state space of this model S has three factors which are the unit quality state, training level, and environmental state, so

$$S = \{(i,k,m) \in S,\ i = 1,2,\cdots,\ N, k = 1,2,\cdots, L \text{ and } m = 1,2,\cdots,, M\}$$

where i, k and m mean the unit quality state, training level and the environmental situation respectively.

Maximum Expected Period, $V(i,k,m)$ When the unit is in quality state i, training level k and the environmental situation is in state m, $V(i,k,m)$ is the maximum expected number of periods until a catastrophic event occurs.

Optimality Equation Because we are looking for the best action policy, the optimality equation selects the best of the three actions.

$$\begin{aligned} V(i,k,m) &= \max\{W_1(i,k,m), W_2(k,m), \delta_{iN} W_3(i,k,m)\} \qquad (5.1) \\ &\quad \text{where} \\ \delta_{iN} &= 0 \text{ if } i = N \\ \delta_{iN} &= 1 \text{ otherwise} \end{aligned}$$

where $W_1(i,k,m)$ is the expected period until a catastrophic event if nothing is done now, $W_2(k,m)$ is the expected period until a catastrophic event if a repair is performed now and $W_3(i,k,m)$ is the expected period until a catastrophic event if training is selected now. Hence $W_1(i,k,m), W_2(k,m), W_3(i,k,m)$ satisfy

$$W_1(i,k,m) = (1-b_m\phi_{iN})(1+\sum_{j=1}^{N}P_{ij}\sum_{k'=1}^{L}T_{kk'}\sum_{m'=1}^{M}S_{mm'}V(j,k',m')) \quad (5.2)$$

where

$$\phi_{iN} = K_k \text{ if } i \neq N$$

$$\phi_{iN} = 1 \text{ if } i = N$$

$$W_2(k,m) = (1-b_m)(1+\sum_{r=1}^{N}R_r\sum_{k'=1}^{L}T_{kk'}\sum_{m'=1}^{M}S_{mm'}V(r,k',m')) \quad (5.3)$$

$$W_3(i,k,m) \quad (5.4)$$

$$= \{1-b_m\times[t+(1-t)K_k]\}(1+\sum_{j=1}^{N}\widetilde{P}_{ij}\sum_{p=1}^{L}w_p\sum_{m'=1}^{M}S_{mm'}V(j,p,m'))$$

(5.2), (5.3), (5.4) can be solved by value iteration where the nth iterate satisfies

$$V_n(i,k,m) = \max\{W_n^1(i,k,m), W_n^2(k,m), \delta_{iN}W_n^3(i,k,m)\} \quad (5.5)$$

where

$$\delta_{iN} = 0 \text{ if } i = N$$

$$\delta_{iN} = 1 \text{ otherwise, and}$$

$$W_n^1(i,k,m) = (1-b_m\delta_{iN})(1+\sum_{j=1}^{N}P_{ij}\sum_{k'=1}^{L}T_{kk'}\sum_{m'=1}^{M}S_{mm'}V_{n-1}(j,k',m')) \quad (5.6)$$

where

$$\phi_{iN} = K_k \text{ if } i \neq N$$

$$\phi_{iN} = 1 \text{ if } i = N$$

$$W_n^2(k,m) = (1-b_m)(1+\sum_{r=1}^{N} R_r \sum_{k'=1}^{L} T_{kk'} \sum_{m'=1}^{M} S_{mm'} V_{n-1}(r,k',m')) \tag{5.7}$$

$$\begin{aligned} & W_n^3(i,k,m) & (5.8) \\ = & \{1 - b_m \times [t + (1-t)K_k]\}(1+\sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V_{n-1}(j,p,m')) \end{aligned}$$

If we define the terminal value, $V_0(i,k,m) = 0$, $V_n(i,k,m)$ is a bounded increasing sequence of function and so converges to the limit $V(i,k,m)$. Standard results from Markov decision processes[Putterman,1994] show that the limit function satisfies the optimality equation (5.1) $\cdots$ (5.4).

Lemma 5.2.1 $V(i,k,m)$ is

a) non-increasing function of i

b) non-increasing function of k

c) non-increasing function of m

where i is the quality state, and k and m are the arbitrary training level and environment situation state.

Proof The proofs use induction hypothesis on n in $V_n(i,k,m)$ and then the result[Putterman,1994] that $V(i,k,m)$ is the limit of the value iteration functions

$V_n(i,k,m)$. Consider a) and defining $V_0(i,k,m) = 0$, then the property holds trivially for $n = 0$. So assume $V_{n-1}(i,k,m)$ is non-increasing in i. This together with the stochastic ordering condition of QSTPM and WTTPM implies

$$\sum_{j=1}^{N} P_{ij} \sum_{k'=1}^{L} T_{kk'} \sum_{m'=1}^{M} S_{mm'} V_{n-1}(j,k',m') > \sum_{j=1}^{N} P_{(i+1)j} \sum_{k'=1}^{L} T_{kk'} \sum_{m'=1}^{M} S_{mm'} V_{n-1}(j,k',m')$$

$$\sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V_{n-1}(j,p,m') > \sum_{j=1}^{N} \widetilde{P}_{(i+1)j} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V_{n-1}(j,p,m')$$

Hence, we can conclude that $W_n^1(i,k,m) \geq W_n^1(i+1,k,m)$ and $W_n^3(i,k,m) \geq W_n^3(i+1,k,m)$. Since $W_n^2(i,k,m) = W_n^2(i+1,k,m)$ from (5.7), it follows $V_n(i,k,m) \geq V_n(i+1,k,m)$. Hence the result holds for $V_n(i,k,m)$ and by convergence the results hold in the limit for $V(i,k,m)$. The proofs of b) and c) follow in a similar way.

Theorem 5.2.1 a) If the unit is down(in state N), it must be repaired immediately.

b) In state (i,k,l), the standby unit is repaired provided $i \geq i^*(k,l)$

Proof In state $i = N$, the only allowed options are do nothing $W_1(N,k,l)$ or repair $W_2(N,k,l) = W_2(k,l)$ since it does not depend on state N.

$$\begin{aligned}
& W_2(k,l) - W_1(N,k,l) \\
= \;& (1-b_l)(1 + \sum_{r=1}^{N} R_r \sum_{k'=1}^{L} T_{kk'} \sum_{m'=1}^{M} S_{lm'} V(r,k',m')) \\
& -(1-b_l)(1 + \sum_{k'=1}^{L} T_{kk'} \sum_{m'=1}^{M} S_{lm'} V(N,k',m'))
\end{aligned}$$

$$= (1-b_l)\sum_{k'=1}^{L} T_{kk'} \sum_{m'=1}^{M} S_{lm'}[\sum_{r=1}^{N} R_r V(r,k',m') - V(N,k',m')]$$

$$= (1-b_l)\sum_{k'=1}^{L} T_{kk'} \sum_{m'=1}^{M} S_{lm'}[\sum_{r=1}^{N} R_r V(r,k',m') - \sum_{r=1}^{N} R_r V(N,k',m')]$$

$$= (1-b_l)\sum_{k'=1}^{L} T_{kk'} \sum_{m'=1}^{M} S_{lm'}\{\sum_{r=1}^{N-1} R_r [V(r,k',m') - V(N,k',m')]\}$$

Since $(1-b_l) \geq 0$, and $V(r,k',m') - V(N,k',m') \geq 0, r < N$, then $W^2(k,l) - W^1(N,k,l) \geq 0$. This means that , at quality state N, we should always repair.

The proof of b) follows because $V(i,k,l)$ is non-increasing function in i and using again stochastic dominance of $P_{ij}, \widetilde{P}_{ij}$ ensures non-increasing property carries through to $\sum_j P_{ij} V(j,k,l)$. Hence $W_1(i,k,l)$ and $W_3(i,k,l)$ are non-increasing in i. Since $W_2(i,k,l)$ is independent of i, once $W_2(i,k,l) \geq W_3(i,k,l)$ and $W_2(i,k,l) \geq W_1(i,k,l)$, then the same inequalities must hold for larger i. So one repairs if $i \geq i^*(k,l)$.

Whereas when the equipment is in its worst state, Theorem 5.2.1 says one needs to repair it immediately, if the training levels are at their worst, it is not always the case that one should train. One needs to add same extra condition as Theorem 5.2.2 implies.

<u>Theorem 5.2.2</u> If $K_L = 1, i = 1, P_{ij} = \widetilde{P}_{ij}$, all i,j and $P_{1\bullet}$ stochastically dominants $R_\bullet$, then one always trains in state $i = 1, k = L$.

<u>Proof</u> If training level is L with $K_L = 1$ and $i = 1$, $(5.2), (5.3), (5.4)$ can be rewritten by

$$W_1(1,L,m) = (1-b_m)(1+\sum_{j=1}^{N} P_{1j} \sum_{m'=1}^{M} S_{mm'} V(j,L,m'))$$
$$W_2(1,L,m) = (1-b_m)(1+\sum_{r=1}^{N} R_r \sum_{m'=1}^{M} S_{mm'} V(r,L,m'))$$
$$W_3(1,L,m) = (1-b_m)(1+\sum_{j=1}^{N} \widetilde{P}_{1j} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V(j,p,m'))$$

Comparing do nothing and training gives

$$\begin{aligned}
& W_3(1,L,m) - W_1(1,L,m) \\
= \; & (1-b_m)(1+\sum_{j=1}^{N} \widetilde{P}_{1j} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V(j,p,m')) \\
& -(1-b_m)(1+\sum_{j=1}^{N} P_{1j} \sum_{m'=1}^{M} S_{mm'} V(j,L,m')) \\
= \; & (1-b_m)[\sum_{j=1}^{N} \widetilde{P}_{1j} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V(j,p,m') - \sum_{j=1}^{N} P_{1j} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V(j,L,m')] \\
= \; & (1-b_m) \sum_{m'=1}^{M} S_{mm'} \sum_{p=1}^{L} w_p [\sum_{j=1}^{N} \widetilde{P}_{1j} V(j,p,m') - \sum_{j=1}^{N} P_{1j} V^{n-1}(j,L,m')] \\
= \; & (1-b_m) \sum_{m'=1}^{M} S_{mm'} \sum_{p=1}^{L} w_p \sum_{j=1}^{N} P_{1j} [V(j,p,m') - V(j,L,m')]
\end{aligned}$$

Since $(1-b_m) \geq 0$ and $V(i,k,m)$ is a non-increasing function in k, $W_3(1,L,m) \geq W_1(1,L,m)$.

$$\begin{aligned}
& W_3(1,L,m) - W_2(L,m) \\
= \; & (1-b_m)(1+\sum_{j=1}^{N} \widetilde{P}_{1j} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V(j,p,m'))
\end{aligned}$$

$$-(1-b_m)(1+\sum_{r=1}^{N} R_r \sum_{m'=1}^{M} S_{mm'} V(r,L,m'))$$

$$= (1-b_m)[\sum_{j=1}^{N} \widetilde{P}_{1j} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V(j,p,m') - \sum_{r=1}^{N} R_r \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V(r,L,m')]$$

$$= (1-b_m) \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} [\sum_{j=1}^{N} \widetilde{P}_{1j} V(j,p,m') - \sum_{r=1}^{N} R_r V(r,L,m')]$$

Since $(1-b_m) \geq 0$, $V(i,k,m)$ is a non-increasing function in k and $\sum_{j=1}^{l} P_{1j} \geq \sum_{r=1}^{l} R_r$(stochastic ordering) where l is an arbitrary quality state, $W_3(1,L,m) \geq W_2(L,m)$. Therefore if the training level L has $K_L = 1$ where $P_{ij} = \widetilde{P}_{ij}$ for all $i, j, \sum_{j=1}^{l} P_{1j} \geq \sum_{r=1}^{l} R_r$ for all l, then training is always optimal.

With slightly weaker condition we can show we also repair or train in the worst training state.

<u>Theorem 5.2.3</u> If $K_L = 1$, and $P_{ij} = \widetilde{P}_{ij}$ $\forall_{ij}$, then in state (i, L, m) one trains if $i < i^*(m)$ and repairs if $i \geq i^*(m)$.

<u>Proof</u> It is enough to show $W_1(i,L,m) \leq \max\{W_2(i,L,m), W_3(i,L,m)\}$ since then the non-increasingness in i of $W_3(i,L,m)$ and the fact $W_2(i,L,m)$ is independent of i gives the rest of the result. Since $P_{ij} = \widetilde{P}_{ij}$ and $\sum w_p V(j,p,m') \geq V(j,L,m')$ by non-increasing property of $V(j,k,m')$ in k, $W_3(i,L,m) \geq W_1(i,L,m)$ for all i and m.

If one considers equipment which is not repairable and so maintenance has no effect, then the only actions possible are training or doing nothing.

The optimality equation for $V(i,k,m)$ in this case satisfies

$$V(i,k,m) = \max\{W_1(i,k,m), \delta_{iN} W_3(i,k,m)\} \quad (5.9)$$

where

$$\delta_{iN} = 0 \text{ if } i = N$$

$$\delta_{iN} = 1 \text{ otherwise}$$

and $W_1(i,k,m)$ and $W_3(i,k,m)$ are still defined by (5.2) and (5.4). In this case one can show that if one decides to train in state (i,k',m) one should train all states $(i,k',m), k' \geq k$.

Theorem 5.2.4 In the non-repairable equipment special case, if one trains in state (i,k,m), one should train in all states $(i,k',m), k' \geq k$.

Proof

$$\begin{aligned} & V(i,k,m) \\ = \ & \max\{W_1(i,k,m), W_3(i,k,m)\} \\ = \ & \max\{(1 - b_m K_k)(1 + \sum_{j=1}^{N} P_{ij} \sum_{k'=1}^{L} T_{kk'} \sum_{m'=1}^{M} S_{mm'} V(j,k',m')), \\ & [1 - b_m \times (t + (1-t)K_k)](1 + \sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V(j,p,m'))\} \end{aligned}$$

If training is optimal at training level k,

$$\begin{aligned} & W_3(i,k,m) \\ = \ & [1 - b_m \times (t + (1-t)K_k)](1 + \sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V(j,p,m')) \end{aligned}$$

$$\begin{aligned} &\geq (1 - b_m K_k)(1 + \sum_{j=1}^{N} P_{ij} \sum_{k'=1}^{L} T_{kk'} \sum_{m'=1}^{M} S_{mm'} V(j, k', m')) \\ &= W_1(i, k, m) \end{aligned}$$

If we let $1 - b_m \times (t + (1-t)K_k) = A(k)$, $1 + \sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{p=1}^{L} w_p \sum_{m'=1}^{M} S_{mm'} V(j, p, m') = B$, $1 - b_m K_k = C(k)$, $1 + \sum_{j=1}^{N} P_{ij} \sum_{k'=1}^{L} T_{kk'} \sum_{m'=1}^{M} S_{mm'} V(j, k', m') = D(k)$, the above relation can be rewritten by

$$\begin{aligned} A(k)B &\geq C(k)D(k) \\ B &\geq \frac{C(k)}{A(k)} D(k) \end{aligned}$$

If we think of $\frac{C(k)}{A(k)}$,

$$\frac{C(k)}{A(k)} = \frac{1 - b_m K_k}{1 - b_m \times (t + (1-t)K_k)}$$

If $b_m, t > 0$, $\frac{C(k)}{A(k)}$ is a decreasing function of k from $\frac{1}{1 - b_m t}$ at $k = 1$(perfect training level : $K_k = 0$) to 1 at $k = L$(worst training level $K_k = 1$).

Because of this and $D(k)$ being a decreasing function in k, hence

$$B \geq \frac{C(k)}{A(k)} D(k) \geq \frac{C(k+1)}{A(k+1)} D(k) \geq \frac{C(k+1)}{A(k+1)} D(k+1)$$

Hence,

$$A(k+1)B \geq C(k+1)D(k+1)$$

This means that $W_3(i, k+1, m) \geq W_1(i, k, m)$ and so if training is optimal at (i, k, m), it is also optimal for (i, k', m), $k' > k$.

5.3 Numerical Example

Consider a problem with five environment situations, 1(most peaceful environment), 2, 3, 4, 5 (most dangerous environment), two training levels $1, 2$, 10 unit quality states, 1(new), $2, \cdots$, 9, 10 (down) and with the Quality State Transition Probability Matrix(QSTPM) and Wear and Tear Transition Probability Matrix(WTTPM) given by Table 5-3 and 5-4. We also assume that repair is not perfect but given by R_r in Table 5-1. The probability of an initiating event, b_m is $\{0.1, 0.2, 0.4, 0.6, 0.7\}$ from environmental situation 1 to 5. The value of K_k, w_p is $K_k = (0,1), w_p = (1,0)$ for training level 1,2. The probability that training can not respond to an initiating event, t is 0.7. We outline variants of this problem with different training transitions. For example 1 $T_{kk'} = \begin{bmatrix} 0.6 & 0.4 \\ 0 & 1 \end{bmatrix}$, for example 2, $T_{kk'} = \begin{bmatrix} 0.01 & 0.99 \\ 0 & 1 \end{bmatrix}$, so training has a positive effect for $\frac{1}{0.4} = 2.5$ periods in example 1 and $\frac{1}{0.99} = 1$ period in example 2.

The results for example 1 and 2 are shown in Table 5-6, 5-7 and Figure 5-1, 5-2, 5-3, 5-4. We can know that the expected survival period is non-increasing function in quality state i, training level k and environment situation m. Because the transition probability from training level 1 to training level 2 in $T_{kk'}$ for example 2 is bigger than in $T_{kk'}$ for example 1, the expected survival period for the example 1 is longer than that for the example 2. Since training is more important for survival in example 2, more training states are selected in training level 1 as we can see in Table 5-7 and Figure 5-3. In Table 5-6 and Figure 5-2, it looks that training is always optimal when $k = L$(i.e. level 2 in example 1) and $i \neq N$. However, this is not true. We can confirm in example 2. From Table 5-7 and Figure 5-4, repair is optimal even though training level is $k = 2$(worst training level) and quality state is in working condition($i = 9$).

Example 3 is the case of $P_{ij} = \widetilde{P}_{ij}$. In reality this means that training does not cause any more wear and tear than in doing nothing case. In QSTPM of this example, $P_{i,i+1} = \widetilde{P}_{i,i+1} = 1$ where $i = 1, \cdots, 5$.Otherwise, $P_{iN} = \widetilde{P}_{iN} = 1$ where $i = 6, - - -, N$. We use the $T_{kk'}$ used in example 2. The other conditions are the same as in previous examples. The result for example 3 is shown in Table 5-8 and Figure 5-5, 5-6. Training is always optimal when $k = 2$(worst training level) and $i = 1$ and we confirm that optimal policy is training, then repair along quality state i when $k = 2$(worst training level, L), $K_2 = 1$ and $P_{ij} = \widetilde{P}_{ij}$.

Example 4 looks only at the do nothing or training problem where repair is not possible. We assume 5 training levels for this example and the TLTPM is in Table 5-2. The values of K_k and w_p are $K_k = (0.0, 0.2, 0.5, 0.7, 1.0)$, $w_p = (0.6, 0.2, 0.1, 0.05, 0.05)$ for training level $k = 1, 2, 3, 4, 5$ respectively. The QSTPM, WTTPM, and t are the same as in example 1 and 2. Table 5-9 and Figure 5-7, 5-8 show the result of the example 4. Hence once training is optimal at training level k, training always optimal for all $k' \geq k$.

Table 5-1. Repair TPM, R_r

r	1	2	3	4	5	6	7	8	9	10
R_r	0.2	0.2	0.1	0.1	0.1	0.1	0.08	0.05	0.05	0.02

Table 5-2. Training Level TPM for example 4, $T_{kk'}$

$k \backslash k'$	1	2	3	4	5
1	0.3	0.3	0.2	0.1	0.1
2	0	0.3	0.3	0.2	0.2
3	0	0	0.4	0.3	0.3
4	0	0	0	0.4	0.6
5	0	0	0	0	1

Table 5-3. Quality State TPM, P_{ij}

$i \setminus j$	1	2	3	4	5	6	7	8	9	10
1	0.2	0.2	0.2	0.1	0.08	0.05	0.05	0.05	0.05	0.02
2	0	0.2	0.2	0.2	0.1	0.1	0.08	0.05	0.04	0.03
3	0	0	0.2	0.2	0.2	0.1	0.1	0.1	0.05	0.05
4	0	0	0	0.2	0.2	0.2	0.15	0.1	0.1	0.05
5	0	0	0	0	0.2	0.3	0.2	0.1	0.1	0.1
6	0	0	0	0	0	0.2	0.3	0.2	0.2	0.1
7	0	0	0	0	0	0	0.2	0.3	0.3	0.2
8	0	0	0	0	0	0	0	0.3	0.4	0.3
9	0	0	0	0	0	0	0	0	0.4	0.6
10	0	0	0	0	0	0	0	0	0	1

Table 5-4. Environmental Situation TPM, $S_{mm'}$

$m \setminus m'$	1	2	3	4	5
1	0.4	0.3	0.2	0.05	0.05
2	0.2	0.4	0.23	0.1	0.07
3	0.1	0.2	0.4	0.2	0.1
4	0.05	0.15	0.2	0.3	0.3
5	0.05	0.1	0.15	0.2	0.5

Table 5-5. Wear and Tear TPM, $\widetilde{P}_{ij}$

$i \setminus j$	1	2	3	4	5	6	7	8	9	10
1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.05	0.05
2	0	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1
3	0	0	0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.1
4	0	0	0	0.1	0.1	0.2	0.2	0.2	0.1	0.1
5	0	0	0	0	0.1	0.1	0.2	0.2	0.2	0.2
6	0	0	0	0	0	0.1	0.2	0.3	0.2	0.2
7	0	0	0	0	0	0	0.1	0.3	0.3	0.3
8	0	0	0	0	0	0	0	0.2	0.4	0.4
9	0	0	0	0	0	0	0	0	0.3	0.7
10	0	0	0	0	0	0	0	0	0	1

Table 5-6. Expected survival time under different actions and choice of optimal action for example 1

m	$k \backslash i$		1	2	3	4	5	6	7	8	9	10
1	1	$W_1(DN)$	6.16	6.08	5.99	5.91	5.83	5.73	5.60	5.51	5.36	4.69
		$W_2(R)$	5.51	5.51	5.51	5.51	5.51	5.51	5.51	5.51	5.51	5.51
		$W_3(T)$	5.99	5.88	5.83	5.76	5.61	5.58	5.47	5.38	5.21	–
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$
	2	$W_1(DN)$	4.98	4.92	4.85	4.78	4.72	4.65	4.56	4.50	4.42	4.38
		$W_2(R)$	4.95	4.95	4.95	4.95	4.95	4.95	4.95	4.95	4.95	4.95
		$W_3(T)$	5.80	5.69	5.65	5.57	5.43	5.40	5.29	5.21	5.04	–
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

m	$k \backslash i$		1	2	3	4	5	6	7	8	9	10
2	1	$W_1(DN)$	5.92	5.84	5.74	5.66	5.57	5.47	5.31	5.20	5.00	3.85
		$W_2(R)$	4.71	4.71	4.71	4.71	4.71	4.71	4.71	4.71	4.71	4.71
		$W_3(T)$	5.41	5.29	5.25	5.18	5.01	4.98	4.85	4.74	4.52	–
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	4.11	4.06	4.00	3.94	3.89	3.82	3.74	3.68	3.61	3.57
		$W_2(R)$	4.09	4.09	4.09	4.09	4.09	4.09	4.09	4.09	4.09	4.09
		$W_3(T)$	5.04	4.92	4.88	4.82	4.66	4.63	4.51	4.41	4.21	–
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

m	$k\backslash i$		1	2	3	4	5	6	7	8	9	10
3	1	$W_1(DN)$	5.51	5.43	5.33	5.24	5.13	5.02	4.83	4.69	4.41	2.47
		$W_2(R)$	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29
		$W_3(T)$	4.35	4.23	4.19	4.13	3.95	3.92	3.77	3.64	3.37	–
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	2.67	2.64	2.59	2.55	2.51	2.47	2.40	2.36	2.30	2.27
		$W_2(R)$	2.66	2.66	2.66	2.66	2.66	2.66	2.66	2.66	2.66	2.66
		$W_3(T)$	3.63	3.52	3.49	3.44	3.29	3.26	3.14	3.03	2.81	–
		Opt.Act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

m	$k\backslash i$		1	2	3	4	5	6	7	8	9	10
4	1	$W_1(DN)$	5.10	5.02	4.91	4.82	4.70	4.58	4.37	4.20	3.84	1.39
		$W_2(R)$	2.03	2.03	2.03	2.03	2.03	2.03	2.03	2.03	2.03	2.03
		$W_3(T)$	3.36	3.24	3.21	3.16	2.98	2.95	2.80	2.68	2.39	–
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	1.51	1.49	1.46	1.44	1.42	1.39	1.35	1.32	1.28	1.26
		$W_2(R)$	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50
		$W_3(T)$	2.31	2.24	2.21	2.18	2.05	2.04	1.93	1.84	1.64	–
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

m	$k\backslash i$		1	2	3	4	5	6	7	8	9	10
5	1	$W_1(DN)$	4.91	4.83	4.72	4.63	4.50	4.38	4.16	3.97	3.58	0.95
		$W_2(R)$	1.46	1.46	1.46	1.46	1.46	1.46	1.46	1.46	1.46	1.46
		$W_3(T)$	2.89	2.79	2.76	2.71	2.54	2.52	2.37	2.25	1.97	–
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	1.04	1.02	1.01	0.99	0.97	0.95	0.92	0.90	0.87	0.86
		$W_2(R)$	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03
		$W_3(T)$	1.70	1.64	1.62	1.59	1.49	1.48	1.39	1.32	1.15	–
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

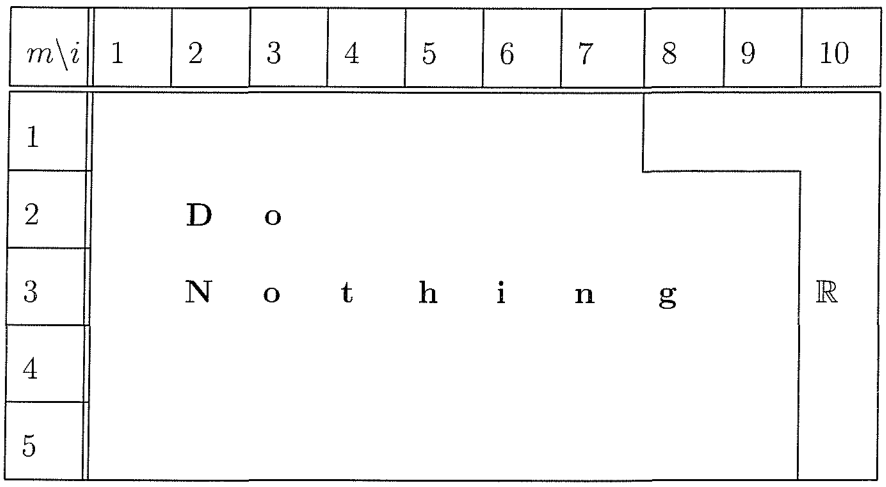

Figure 5-1. Result of example 1, training level 1(simple form)

<table>
<tr><th>$m\backslash i$</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th></tr>
<tr><td>1</td><td colspan="9" rowspan="5">T</td><td rowspan="5">$\mathbb{R}$</td></tr>
<tr><td>2</td></tr>
<tr><td>3</td></tr>
<tr><td>4</td></tr>
<tr><td>5</td></tr>
</table>

Figure 5-2. Result of example 1, training level 2(simple form)

Table 5-7. Expected survival time under different actions and choice of optimal action for example 2

m		$k\backslash\ i$	1	2	3	4	5	6	7	8	9	10
1	1	$W_1(DN)$	4.43	4.40	4.36	4.32	4.28	4.23	4.17	4.14	4.11	3.69
		$W_2(R)$	3.98	3.98	3.98	3.98	3.98	3.98	3.98	3.98	3.98	3.98
		$W_3(T)$	4.53	4.46	4.44	4.41	4.31	4.30	4.22	4.15	3.98	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	3.98	3.95	3.92	3.88	3.85	3.80	3.75	3.72	3.70	3.69
		$W_2(R)$	3.97	3.97	3.97	3.97	3.97	3.97	3.97	3.97	3.97	3.97
		$W_3(T)$	4.38	4.32	4.30	4.26	4.17	4.16	4.08	4.02	3.85	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$

m	$k\backslash\ i$		1	2	3	4	5	6	7	8	9	10
2	1	$W_1(DN)$	4.12	4.09	4.05	4.01	3.97	3.92	3.86	3.82	3.78	3.02
		$W_2(R)$	3.28	3.28	3.28	3.28	3.28	3.28	3.28	3.28	3.28	3.28
		$W_3(T)$	4.03	3.97	3.95	3.92	3.82	3.81	3.72	3.65	3.45	-
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	3.29	3.27	3.23	3.20	3.17	3.13	3.08	3.05	3.03	3.02
		$W_2(R)$	3.28	3.28	3.28	3.28	3.28	3.28	3.28	3.28	3.28	3.28
		$W_3(T)$	3.75	3.69	3.67	3.65	3.56	3.54	3.46	3.39	3.21	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$

m	$k\backslash\ i$		1	2	3	4	5	6	7	8	9	10
3	1	$W_1(DN)$	3.59	3.56	3.52	3.49	3.45	3.40	3.34	3.29	3.24	1.94
		$W_2(R)$	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15	2.15
		$W_3(T)$	3.17	3.11	3.10	3.07	2.97	2.96	2.87	2.78	2.55	-
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	2.15	2.13	2.11	2.08	2.06	2.03	1.99	1.97	1.95	1.94
		$W_2(R)$	2.14	2.14	2.14	2.14	2.14	2.14	2.14	2.14	2.14	2.14
		$W_3(T)$	2.64	2.59	2.58	2.56	2.47	2.46	2.39	2.32	2.13	-
		Opt.Act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$

m	$k\backslash i$		1	2	3	4	5	6	7	8	9	10
4	1	$W_1(DN)$	3.08	3.05	3.02	2.98	2.95	2.90	2.84	2.80	2.75	1.09
		$W_2(R)$	1.22	1.22	1.22	1.22	1.22	1.22	1.22	1.22	1.22	1.22
		$W_3(T)$	2.39	2.34	2.33	2.31	2.21	2.20	2.11	2.03	1.81	-
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	1.23	1.21	1.20	1.19	1.17	1.15	1.13	1.11	1.10	1.09
		$W_2(R)$	1.22	1.22	1.22	1.22	1.22	1.22	1.22	1.22	1.22	1.22
		$W_3(T)$	1.65	1.61	1.60	1.59	1.52	1.52	1.46	1.40	1.25	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

m	$k\backslash i$		1	2	3	4	5	6	7	8	9	10
5	1	$W_1(DN)$	2.84	2.82	2.78	2.75	2.72	2.68	2.62	2.58	2.53	0.75
		$W_2(R)$	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
		$W_3(T)$	2.04	1.99	1.98	1.96	1.87	1.86	1.78	1.70	1.49	-
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	0.85	0.84	0.83	0.82	0.81	0.80	0.78	0.77	0.76	0.75
		$W_2(R)$	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
		$W_3(T)$	1.20	1.17	1.16	1.15	1.10	1.09	1.05	1.00	0.87	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

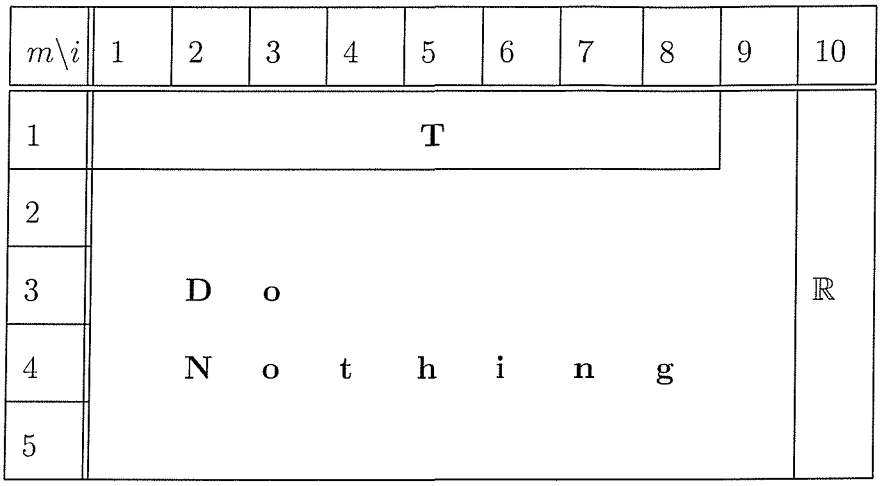

Figure 5-3. Result of example 2, training level 1(simple form)

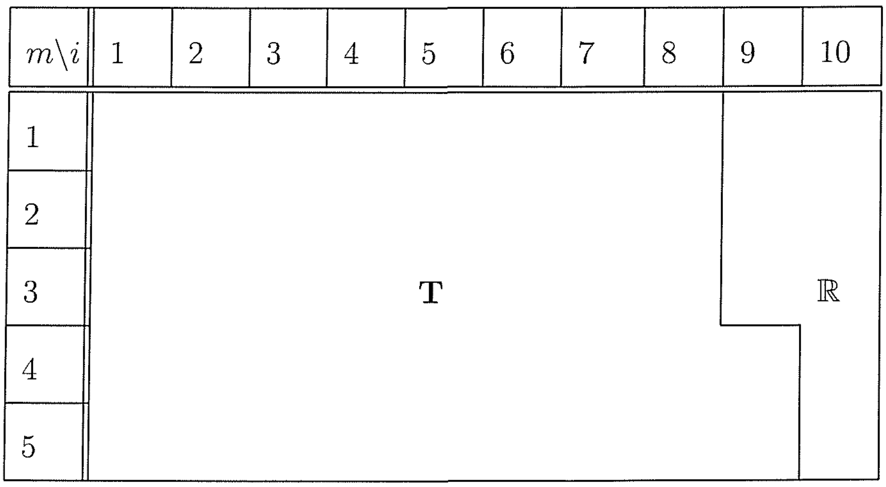

Figure 5-4. Result of example 2, training level 2(simple form)

Table 5-8. Expected survival time under different actions and choice of optimal action for example 3

m	$k\backslash i$		1	2	3	4	5	6	7	8	9	10
1	1	$W_1(DN)$	4.81	4.74	4.57	4.49	4.23	4.22	4.22	4.22	4.22	3.80
		$W_2(R)$	4.13	4.13	4.13	4.13	4.13	4.13	4.13	4.13	4.13	4.13
		$W_3(T)$	4.98	4.86	4.76	4.56	4.45	3.93	3.93	3.93	3.93	-
		act	**T**	**T**	**T**	**T**	**T**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	4.33	4.27	4.11	4.04	3.80	3.80	3.80	3.80	3.80	3.80
		$W_2(R)$	4.13	4.13	4.13	4.13	4.13	4.13	4.13	4.13	4.13	4.13
		$W_3(T)$	4.82	4.70	4.60	4.41	4.30	3.81	3.81	3.81	3.81	-
		act	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$

m	$k\backslash i$		1	2	3	4	5	6	7	8	9	10
2	1	$W_1(DN)$	4.48	4.42	4.25	4.19	3.89	3.88	3.88	3.88	3.88	3.10
		$W_2(R)$	3.41	3.41	3.41	3.41	3.41	3.41	3.41	3.41	3.41	3.41
		$W_3(T)$	4.45	4.32	4.26	4.04	3.97	3.34	3.34	3.34	3.34	-
		act	**DN**	**DN**	**T**	**DN**	**T**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	3.57	3.53	3.39	3.34	3.10	3.10	3.10	3.10	3.10	3.10
		$W_2(R)$	3.41	3.41	3.41	3.41	3.41	3.41	3.41	3.41	3.41	3.41
		$W_3(T)$	4.14	4.02	3.96	3.76	3.69	3.11	3.11	3.11	3.11	-
		act	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$

m	$k\backslash i$		1	2	3	4	5	6	7	8	9	10
3	1	$W_1(DN)$	3.90	3.86	3.70	3.66	3.33	3.32	3.32	3.32	3.32	1.99
		$W_2(R)$	2.22	2.22	2.22	2.22	2.22	2.22	2.22	2.22	2.22	2.22
		$W_3(T)$	3.51	3.41	3.36	3.16	3.12	2.39	2.39	2.39	2.39	-
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	2.33	2.31	2.21	2.19	1.99	1.99	1.99	1.99	1.99	1.99
		$W_2(R)$	2.21	2.21	2.21	2.21	2.21	2.21	2.21	2.21	2.21	2.21
		$W_3(T)$	2.93	2.84	2.80	2.63	2.60	2.00	2.00	2.00	2.00	-
		Opt.Act	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$

m	$k\backslash i$		1	2	3	4	5	6	7	8	9	10
4	1	$W_1(DN)$	3.34	3.31	3.17	3.14	2.81	2.79	2.79	2.79	2.79	1.11
		$W_2(R)$	1.26	1.26	1.26	1.26	1.26	1.26	1.26	1.26	1.26	1.26
		$W_3(T)$	2.66	2.58	2.55	2.37	2.35	1.62	1.62	1.62	1.62	-
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	1.33	1.31	1.26	1.25	1.11	1.11	1.11	1.11	1.11	1.11
		$W_2(R)$	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25
		$W_3(T)$	1.83	1.77	1.76	1.64	1.62	1.12	1.12	1.12	1.12	-
		act	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$

m	$k\backslash i$		1	2	3	4	5	6	7	8	9	10
5	1	$W_1(DN)$	3.08	3.05	2.93	2.90	2.57	2.56	2.56	2.56	2.56	0.76
		$W_2(R)$	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87
		$W_3(T)$	2.27	2.20	2.18	2.02	2.00	1.31	1.31	1.31	1.31	-
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$
	2	$W_1(DN)$	0.92	0.91	0.87	0.86	0.76	0.76	0.76	0.76	0.76	0.76
		$W_2(R)$	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86
		$W_3(T)$	1.33	1.29	1.28	1.19	1.18	0.77	0.77	0.77	0.77	-
		act	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$

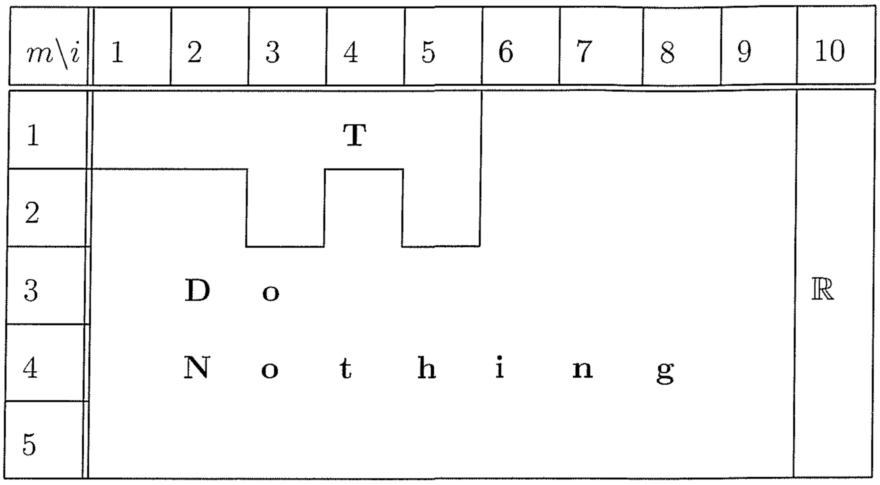

Figure 5-5. Result of example 3, training level 1(simple form)

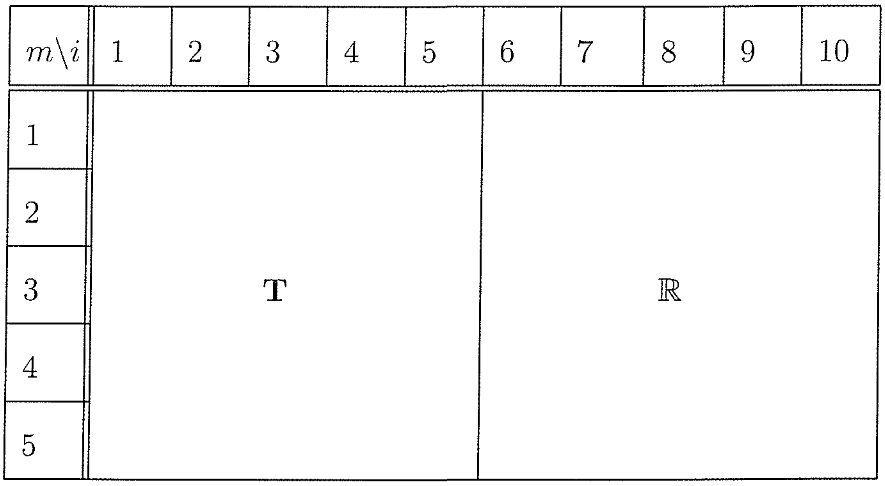

Figure 5-6. Result of example 3, training level 2(simple form)

Table 5-9. Expected survival time under different actions and choice of optimal action for example 4

m	$k\backslash i$		1	2	3	4	5	6	7	8	9	10
1	1	$W_1(DN)$	4.72	4.65	4.54	4.44	4.34	4.22	4.04	3.90	3.62	3.02
		$W_3(T)$	4.45	4.30	4.26	4.18	3.97	3.93	3.77	3.63	3.35	-
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**
	2	$W_1(DN)$	4.38	4.31	4.21	4.13	4.04	3.94	3.79	3.67	3.47	3.02
		$W_3(T)$	4.42	4.28	4.23	4.15	3.95	3.91	3.74	3.60	3.32	-
		act	**T**	**DN**	**T**	**T**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**
	3	$W_1(DN)$	4.13	4.07	3.97	3.89	3.81	3.71	3.58	3.48	3.32	3.02
		$W_3(T)$	4.38	4.24	4.19	4.11	3.91	3.87	3.70	3.57	3.29	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**DN**	**DN**
	4	$W_1(DN)$	3.98	3.92	3.83	3.75	3.66	3.57	3.44	3.35	3.21	3.02
		$W_3(T)$	4.35	4.21	4.16	4.08	3.88	3.85	3.68	3.54	3.27	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**DN**
	5	$W_1(DN)$	3.82	3.76	3.67	3.59	3.52	3.43	3.30	3.22	3.09	3.02
		$W_3(T)$	4.30	4.17	4.12	4.04	3.84	3.81	3.64	3.51	3.24	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**DN**

m	$k \backslash i$		1	2	3	4	5	6	7	8	9	10
5	1	$W_1(DN)$	3.85	3.78	3.68	3.60	3.48	3.38	3.16	2.98	2.58	0.64
		$W_3(T)$	2.09	2.00	1.97	1.93	1.79	1.77	1.65	1.54	1.30	-
		Opt.Act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**
	2	$W_1(DN)$	2.87	2.82	2.76	2.70	2.63	2.57	2.43	2.33	2.10	0.64
		$W_3(T)$	1.91	1.83	1.81	1.77	1.64	1.62	1.51	1.41	1.19	-
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**
	3	$W_1(DN)$	1.98	1.95	1.90	1.87	1.82	1.78	1.70	1.64	1.52	0.64
		$W_3(T)$	1.65	1.58	1.57	1.53	1.42	1.40	1.31	1.22	1.03	-
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**
	4	$W_1(DN)$	1.45	1.43	1.39	1.37	1.33	1.30	1.25	1.21	1.15	0.64
		$W_3(T)$	1.48	1.42	1.40	1.37	1.27	1.26	1.17	1.09	0.92	-
		act	**T**	**DN**	**T**	**T**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**
	5	$W_1(DN)$	0.83	0.81	0.79	0.78	0.76	0.74	0.71	0.69	0.66	0.64
		$W_3(T)$	1.22	1.17	1.16	1.13	1.05	1.04	0.97	0.90	0.76	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**DN**

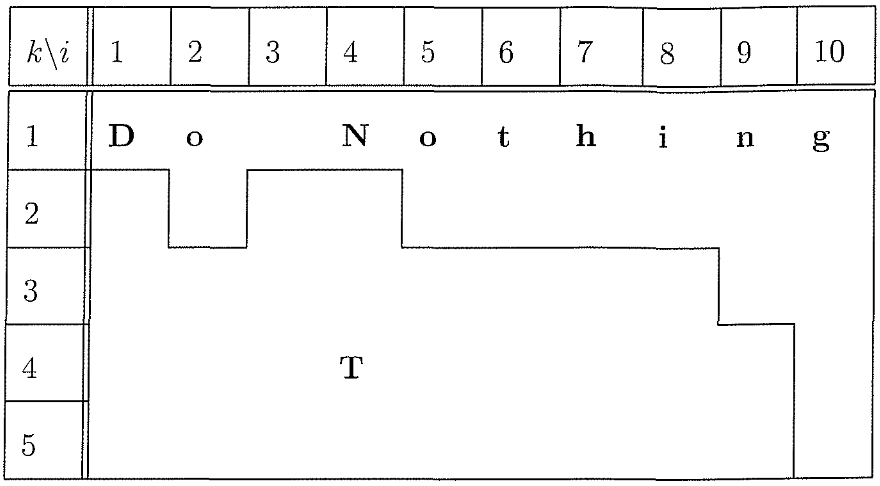

Figure 5-7. Result of example 4, environment 1(simple form)

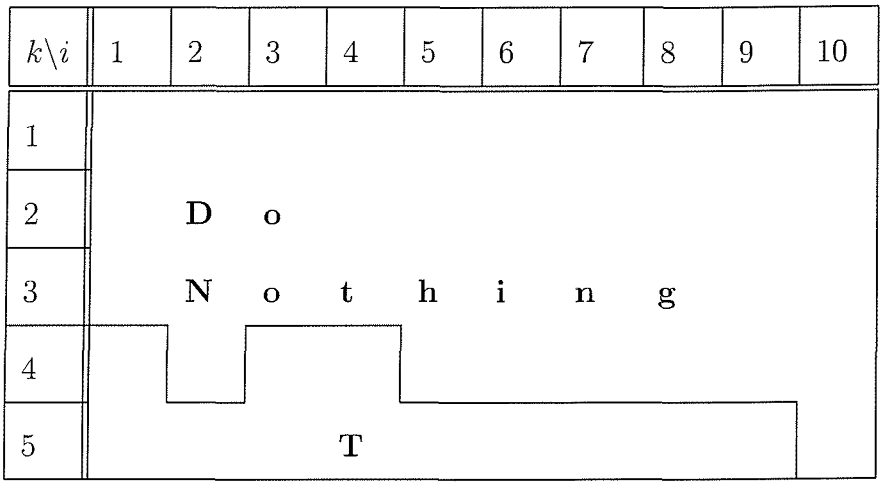

Figure 5-8. Result of example 4, environment 5(simple form)

5.4 Training Model with Continuous Loss of Expertise

5.4.1 Introduction

In this section we consider the problem where the expertise obtained by training on the equipment is gradually lost over time rather than subject to random changes as in section 2. To do this we define an expertise index which shows how well trained the operator of the stand-by unit has been. An operator with higher values in this index is likely to perform better. Apart from the modification of an expertise index instead of a training level, the other conditions are the same as in section 5.2.

5.4.2 Terminology

Expertise Index, T To define the expertise index, we assume that if it is at level $\mathbb{T}, \mathbb{T} \geq 0$, then

a) when no training occurs at the next period, it moves to $\alpha\mathbb{T}$

b) when training occurs at the next period, it moves to $\alpha\mathbb{T} + 1$

So in a sense, expertise obtained through training dissipates geometrically(the equivalent of exponentially in discrete time) and each period of training adds 1 unit to the expertise level whatever it is. Thus training all the time gives us a level of $1 + \alpha + \alpha^2 + \alpha^3 + \cdots = (1-\alpha)^{-1}$, while no training gives an expertise of 0. In order to make the index easy to understand, we multiply the above index by $(1-\alpha)$ to arrive at one where all values are between 0 and 1, and if we let $(1-\alpha)\mathbb{T}, T$, training changes T into $\alpha T + (1-\alpha)$, while no training changes T into αT. If the expertise index is T, the probability the operator can not respond to satisfactorily to an initiating event is f_T where $0 \leq f_T \leq 1$, $f_{T'} < f_T$ if $T < T'$.

5.4.3 Model

Model State Space The state space of this model S has three factors which are the unit quality state, training level, and environmental state, so

$$S = \{(i, T, m) \in S,\ i = 1, 2, ---,\ 0 \leq T \leq 1 \text{ and } m = 1, 2, ---,, M\}$$

where i, T and m mean the unit quality state, expertise index and the environmental situation respectively.

Maximum Expected Period, $V(i,T,m)$ When the unit is in quality state i, expertise index T and the environmental situation is in state m, $V(i,T,m)$ is the maximum expected number of periods until a catastrophic event occurs.

Optimality Equation

$$V(i,T,m) = \max\{W_1(i,T,m), W_2(T,m), \delta_{iN}W_3(i,T,m)\} \tag{5.10}$$

where

$$\delta_{iN} = 0 \text{ if } i = N$$

$$\delta_{iN} = 1 \text{ otherwise}$$

where

$$W_1(i,T,m) = (1 - b_m\delta_{iN})(1 + \sum_{j=1}^{N} P_{ij} \sum_{m'=1}^{M} S_{mm'} V(j, \alpha T, m')) \tag{5.11}$$

where

$$\phi_{iN} = f_T \text{ if } i \neq N$$

$$\phi_{iN} = 1 \text{ if } i = N$$

$$W_2(T,m) = (1 - b_m)(1 + \sum_{r=1}^{N} R_r \sum_{m'=1}^{M} S_{mm'} V(r, \alpha T, m')) \tag{5.12}$$

$$W_3(i,T,m) \tag{5.13}$$
$$= \{1 - b_m \times [t + (1-t)f_T]\}(1 + \sum_{j=1}^{N} \widetilde{P}_{ij} \sum_{m'=1}^{M} S_{mm'} V(j, \alpha T + (1-\alpha), m'))$$

(5.9), (5.10), (5.11), (5.12) can be solved using value iteration. The results of the previous section extend to this model and the proofs follow by induction on value iteration. The value iteration scheme satisfies equation (5.9), (5.10), (5.11), (5.12) with V_n, W_n on the L.H.S. and V_{n-1} on the R.H.S.

Lemma 5.4.1 $V(i,T,m)$ is

a) non-increasing function of i

b) non-increasing function of T

c) non-increasing function of m

where i is the quality state, T and m are the arbitrary expertise index and environment state.

Proof As with lemma 5.2.1, the proofs use induction hypothesis on n in $V_n(i,T,m)$ and then the result that $V(i,T,m)$ is limit of the value iteration functions $V_n(i,T,m)$. If we consider a) and defining $V_0(i,T,m) = 0$, then the property holds trivially for $n = 0$. So assume $V_{n-1}(i,T,m)$ is non-increasing in i. This together with the stochastic ordering condition of P_{ij} and $\widetilde{P}_{ij}$ implies

$$W_n^1(i,T,m) \geq W_n^1(i+1,T,m)$$

$$W_n^3(i,T,m) \geq W_n^3(i+1,T,m)$$

Since $W_n^2(i,T,m) = W_n^2(i+1,T,m)$ from (11), it follows $V_n(i,T,m) \geq V_n(i+1,T,m)$. By convergence the results hold in the limit for $V(i,T,m)$. The proofs of b) and c) follow in a similar way.

Theorem 5.4.1 In state N(down), one should always repair.

Proof Similarly with theorem 5.2.1, in state $i = N$, the only allowed options are do nothing $W_1(N,T,l)$ or repair $W_2(N,T,l) = W_2(T,l)$ since it does not depend on state N.

$$\begin{aligned}
& W_2(T,l) - W_1(N,T,l) \\
= \; & (1-b_l)(1+\sum_{r=1}^{N} R_r \sum_{m'=1}^{M} S_{lm'} V(r,\alpha T,m')) \\
& -(1-b_l)(1+\sum_{m'=1}^{M} S_{lm'} V(N,\alpha T,m')) \\
= \; & (1-b_l)\sum_{m'=1}^{M} S_{lm'} [\sum_{r=1}^{N} R_r V(r,\alpha T,m') - V(N,\alpha T,m')] \\
= \; & (1-b_l)\sum_{m'=1}^{M} S_{lm'} [\sum_{r=1}^{N} R_r V(r,\alpha T,m') - \sum_{r=1}^{N} R_r V(N,\alpha T,m')] \\
= \; & (1-b_l)\sum_{m'=1}^{M} S_{lm'} \{\sum_{r=1}^{N-1} R_r [V(r,\alpha T,m') - V(N,\alpha T,m')]\}
\end{aligned}$$

Since $(1-b_l) \geq 0$, and $V(r,\alpha T,m') - V(N,\alpha T,m') \geq 0, r < N$, then $W_2(T,l) - W_1(N,T,l) \geq 0$. This means that , at quality state N, we should always repair.

5.5 Numerical Example

In this example, there are also 5 different environmental situation states and 10 different unit quality states. The probability of an initiating event, b_m, where $1 \leq m \leq 5$ is $(0.1, 0.2, 0.4, 0.6, 0.7)$. The transition matrices for quality state, environment situation are the same as in previous examples. The probability which training can not respond an initiating event, t is 0.7. The discount factor for the expertness index, α is 0.6.The effect of the repairs has the following distributions in Table 5-9.

Table 5-10. Repair TPM, R_r

r	1	2	3	4	5	6	7	8	9	10
R_r	0.1	0.1	0.2	0.2	0.1	0.1	0.08	0.05	0.05	0.02

The results in Table 5-10 and Figure 5-9 show that the expected survival period is also non-increasing function in quality state and environment situation and non-decreasing in expertness index. The results in general show a pattern in that as the quality state increases(gets worse) one initially trains, then does nothing, and then repair. However this pattern is violated in the case of $T = 0.7$. On the other hand, if the unit is operating then as the expertise index increases, it shows a trend in which one moves from training to doing nothing but this need not always be the case. The results in Table 5-11 and Figure 5-10 under different initiating event probabilities, QSTPM, TLTPM show a movement from training directly to repair.

Table 5-11. The result of example when $b_1 = 0.1, b_2 = 0.2, b_3 = 0.4, b_4 = 0.6, b_5 = 0.7,$

$t = 0.7, \alpha = 0.6$

m	$T \backslash i$		1	2	3	4	5	6	7	8	9	10
3	0.0	$W_1(DN)$	1.88	1.87	1.85	1.84	1.83	1.81	1.79	1.77	1.76	1.75
		$W_2(R)$	1.87	1.87	1.87	1.87	1.87	1.87	1.87	1.87	1.87	1.87
		$W_3(T)$	2.19	2.16	2.14	2.13	2.07	2.06	2.02	1.99	1.91	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

m	$T \backslash i$		1	2	3	4	5	6	7	8	9	10
3	0.2	$W_1(DN)$	2.24	2.23	2.20	2.18	2.16	2.14	2.11	2.09	2.05	1.78
		$W_2(R)$	1.97	1.97	1.97	1.97	1.97	1.97	1.97	1.97	1.97	1.97
		$W_3(T)$	2.50	2.45	2.43	2.41	2.34	2.33	2.28	2.23	2.13	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

m	$T \backslash i$		1	2	3	4	5	6	7	8	9	10
3	0.5	$W_1(DN)$	2.85	2.82	2.79	2.76	2.73	2.70	2.64	2.60	2.52	1.83
		$W_2(R)$	2.12	2.12	2.12	2.12	2.12	2.12	2.12	2.12	2.12	2.12
		$W_3(T)$	3.01	2.94	2.92	2.89	2.80	2.78	2.70	2.63	2.48	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**DN**	$\mathbb{R}$

m	$T\backslash i$		1	2	3	4	5	6	7	8	9	10
3	0.6	$W_1(DN)$	3.07	3.04	3.01	2.97	2.94	2.90	2.83	2.78	2.68	1.84
		$W_2(R)$	2.18	2.18	2.18	2.18	2.18	2.18	2.18	2.18	2.18	2.18
		$W_3(T)$	3.20	3.13	3.11	3.07	2.96	2.95	2.86	2.78	2.60	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**DN**	**DN**	$\mathbb{R}$

m	$T\backslash i$		1	2	3	4	5	6	7	8	9	10
3	0.7	$W_1(DN)$	3.3	3.27	3.23	3.19	3.15	3.11	3.03	2.97	2.85	1.86
		$W_2(R)$	2.23	2.23	2.23	2.23	2.23	2.23	2.23	2.23	2.23	2.23
		$W_3(T)$	3.4	3.32	3.3	3.26	3.14	3.12	3.02	2.93	2.72	-
		act	**T**	**T**	**T**	**T**	**DN**	**T**	**DN**	**DN**	**DN**	$\mathbb{R}$

m	$T\backslash i$		1	2	3	4	5	6	7	8	9	10
3	0.8	$W_1(DN)$	3.54	3.51	3.46	3.42	3.37	3.33	3.24	3.17	3.03	1.88
		$W_2(R)$	2.29	2.29	2.29	2.29	2.29	2.29	2.29	2.29	2.29	2.29
		$W_3(T)$	3.60	3.52	3.49	3.45	3.32	3.30	3.18	3.08	2.85	-
		Opt.Act	**T**	**T**	**T**	**T**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$

m	$T\backslash i$		1	2	3	4	5	6	7	8	9	10
3	0.9	$W_1(DN)$	3.80	3.76	3.71	3.67	3.61	3.56	3.46	3.38	3.22	1.90
		$W_2(R)$	2.36	2.36	2.36	2.36	2.36	2.36	2.36	2.36	2.36	2.36
		$W_3(T)$	3.82	3.73	3.70	3.66	3.51	3.49	3.36	3.24	2.99	-
		act	**T**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$

m	$T\backslash\ i$		1	2	3	4	5	6	7	8	9	10
3	1.0	$W_1(DN)$	4.07	4.02	3.97	3.92	3.86	3.80	3.69	3.60	3.41	1.92
		$W_2(R)$	2.42	2.42	2.42	2.42	2.42	2.42	2.42	2.42	2.42	2.42
		$W_3(T)$	4.05	3.95	3.92	3.87	3.70	3.68	3.53	3.41	3.12	-
		act	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	**DN**	$\mathbb{R}$

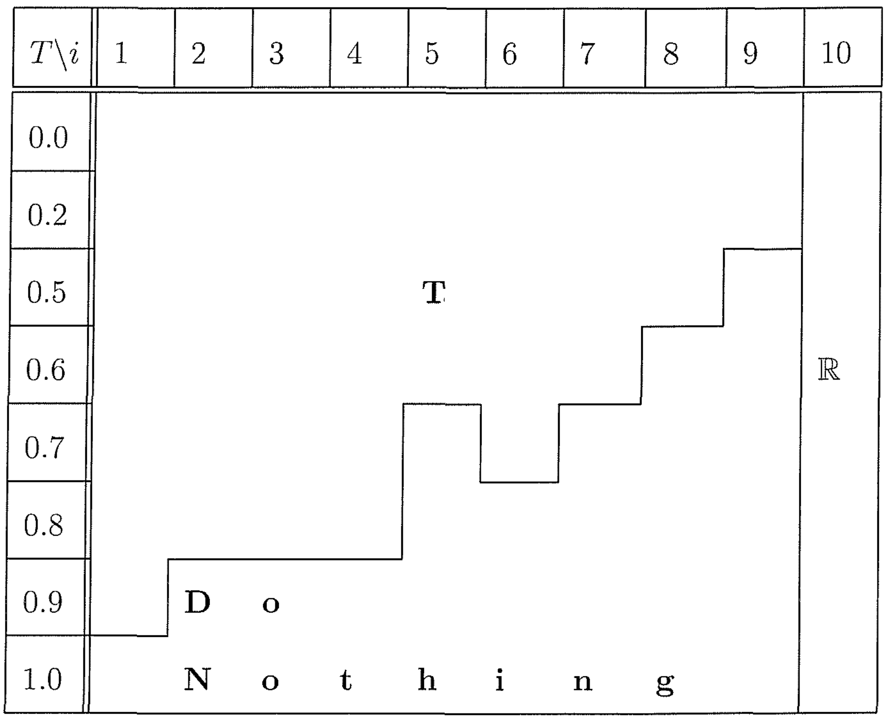

Figure 5-9. Simple form of Table 5-11

Table 5-12. The result of example when $b_1 = 10^{-23}, b_2 = 0.1, b_3 = 0.2, b_4 = 0.3, b_5 = 0.4,$ $t = 0.4, \alpha = 0.6$

m	$T \backslash i$		1	2	3	4	5	6	7	8	9	10
1	0.0	$W_1(DN)$	7.22	7.01	7.01	7.01	7.01	7.01	7.01	7.01	6.92	6.92
		$W_2(R)$	7.06	7.06	7.06	7.06	7.06	7.06	7.06	7.06	7.06	7.06
		$W_3(T)$	7.39	7.34	7.15	7.15	7.15	7.15	7.15	7.15	7.15	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

m	$T \backslash i$		1	2	3	4	5	6	7	8	9	10
1	0.2	$W_1(DN)$	7.34	7.10	7.10	7.10	7.10	7.10	7.10	7.10	6.98	6.98
		$W_2(R)$	7.16	7.16	7.16	7.16	7.16	7.16	7.16	7.16	7.16	7.16
		$W_3(T)$	7.53	7.45	7.24	7.24	7.24	7.24	7.24	7.24	7.24	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

m	$T \backslash i$		1	2	3	4	5	6	7	8	9	10
1	0.5	$W_1(DN)$	7.53	7.28	7.28	7.28	7.28	7.28	7.28	7.25	7.08	7.08
		$W_2(R)$	7.34	7.34	7.34	7.34	7.34	7.34	7.34	7.34	7.34	7.34
		$W_3(T)$	7.76	7.63	7.38	7.38	7.38	7.38	7.38	7.38	7.38	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

m	$T\backslash\ i$		1	2	3	4	5	6	7	8	9	10
1	0.6	$W_1(DN)$	7.59	7.36	7.36	7.36	7.36	7.35	7.34	7.30	7.13	7.13
		$W_2(R)$	7.41	7.41	7.41	7.41	7.41	7.41	7.41	7.41	7.41	7.41
		$W_3(T)$	7.85	7.70	7.43	7.43	7.43	7.43	7.43	7.43	7.43	-
		act	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	**T**	$\mathbb{R}$

m	$T\backslash\ i$		1	2	3	4	5	6	7	8	9	10
3	0.7	$W_1(DN)$	7.66	7.43	7.43	7.43	7.43	7.43	7.41	7.36	7.17	7.17
		$W_2(R)$	7.49	7.49	7.49	7.49	7.49	7.49	7.49	7.49	7.49	7.49
		$W_3(T)$	7.93	7.77	7.48	7.48	7.48	7.48	7.48	7.48	7.48	-
		act	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$

m	$T\backslash\ i$		1	2	3	4	5	6	7	8	9	10
1	0.8	$W_1(DN)$	7.73	7.52	7.52	7.52	7.52	7.51	7.48	7.41	7.21	7.21
		$W_2(R)$	7.56	7.56	7.56	7.56	7.56	7.56	7.56	7.56	7.56	7.56
		$W_3(T)$	8.02	7.84	7.54	7.54	7.54	7.54	7.54	7.54	7.54	7.54
		Opt.Act	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$

m	$T\backslash\ i$		1	2	3	4	5	6	7	8	9	10
1	0.9	$W_1(DN)$	7.80	7.60	7.60	7.60	7.60	7.59	7.56	7.47	7.25	7.25
		$W_2(R)$	7.64	7.64	7.64	7.64	7.64	7.64	7.64	7.64	7.64	7.64
		$W_3(T)$	8.11	7.91	7.59	7.59	7.59	7.59	7.59	7.59	7.59	7.59
		act	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$

m	$T\backslash i$		1	2	3	4	5	6	7	8	9	10
1	1.0	$W_1(DN)$	7.87	7.69	7.69	7.69	7.69	7.68	7.63	7.53	7.30	7.30
		$W_2(R)$	7.73	7.73	7.73	7.73	7.73	7.73	7.73	7.73	7.73	7.73
		$W_3(T)$	8.20	7.99	7.64	7.64	7.64	7.64	7.64	7.64	7.64	7.64
		act	**T**	**T**	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$

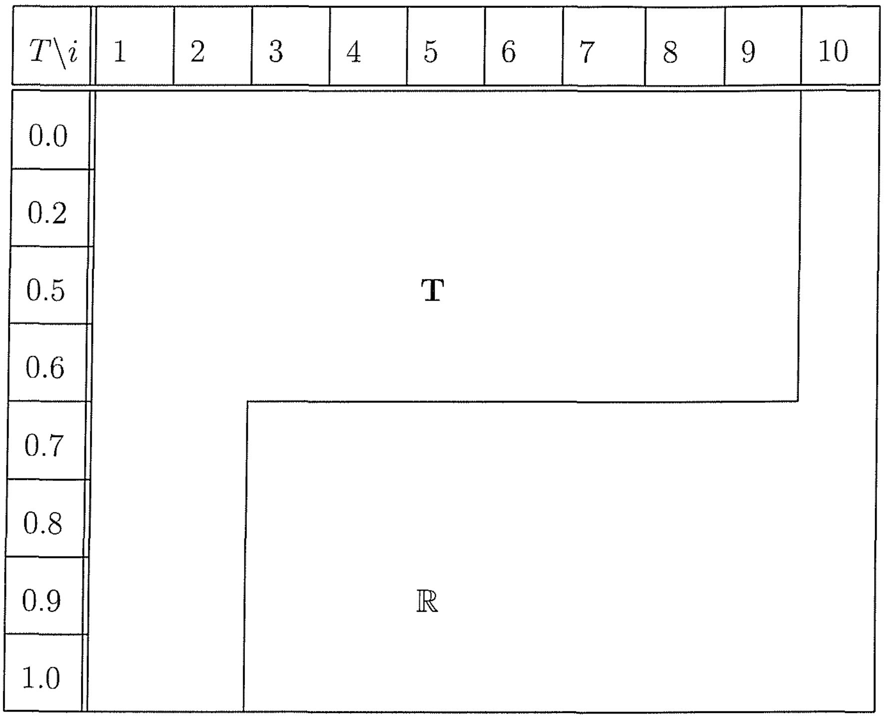

Figure 5-10. Simple form of Table 5-12

5.6 Conclusions

The models presented in this paper show that there can be, depending on parameters, a strong interaction between the quality state of the stand-by unit, the general environment state, the training level of the operator and the decision on whether to repair or train. In the training model with random loss of expertise, the expected survival time until a catastrophic situation decreases as the unit quality states, the training levels and environment situations worsen. One always repairs when the unit is down. Also, once repair is optimal at certain quality state, repair is always optimal for worse quality states. When the unit quality state is new, the training level is at its lowest which means that it can not respond to an initiating event and there is no difference between QSTPM and WTTPM, one always trains or repairs. One trains if the quality state is good and there is same quality level below which one repairs. If one adds the extra condition $\sum_{j<l} P_{1j} \geq \sum_{r<l} R_r$ to the above condition, one can show one always trains when the quality state is new. If the repair action is not available and the quality state is in working condition, once training is optimal at a certain training level, one always trains at worse training levels. In the training model with continuous loss of expertise, we find that one always repairs when the unit is down. If the unit is operating then as the expertise index increases, it shows a trend in which one moves from training to doing nothing but this need not always be the case.

CHAPTER 6

Conclusion

This concluding chapter summarises the main results of the research and proposes future developments. In this thesis, we have dealt with repair strategies of stand-by equipment which maximise the expected survival time until a catastrophic event. A stand-by unit is equipment which is only brought into operation when there is a vital need for it like military equipment, hospital power supply, etc. We have seen that there are many maintenance, repair and replacement models for deteriorating equipment in the literature review of the chapter 2. Almost all the literature in those concentrate on policies which minimise the average discounted cost criterion. Some authors researched the idea of using a catastrophic event criterion to overcome the problem that failure result in unquantifiably large cost. Our research focused on the expected survival time instead because the failure of the standby equipment causes immeasurable cost, i.e. national security, human lives etc. Other authors have looked at maintenance in a random environment but in that case the unit is always in use so the changes in the environment age the equipment at different rates, but do not affect when it is needed. Some papers considered protective systems with non-self-announcing failures where the rate of deterioration is governed by random environment. Our study allowed the deterioration of the equipment to be independent of the environment, but the environment affects the need for the equipment. We also looked at conflict situation where the environment is controlled by an opponent. Finally we developed Markov decision process models which consider training. Equipment can only be used if it is in an operable state and if its users have had sufficient recent training with it. Thus as well as repairing and maintaining the equipment, it is necessary to train users. For this thesis, we developed above three models of repair strategies for stand-by equipment under changing environment. The result of this research is as follows.

6.1 Markov Decision Process approach

In chapter 3, we described the form of optimal repair policy by Markov decision process. Here we considered one repair action and two repair actions cases. The models presented in chapter 3 show that there is a strong interaction between the quality state of the stand-by unit, the general environment situation, and the repair action. In the one repair action case, one always repairs when the unit is down. This is quite obvious because any other actions can not be better than repair when the unit is broken down. Another important finding here is that in the most hostile environment one repairs the equipment only when the unit is down. This means that one needs to do nothing when the unit is working because the probability of the initiating event in this case is the highest. In other environments one can make repairs when the unit is still functioning. This is the case when the repair action for the unit that is in working condition may give longer expected survival time than doing nothing. Generally the states at which one repairs increase, the more benign the environment becomes, but this need not always be the case. We have counter examples.

With quick and slow repairs in two repair action model, the repair or do nothing decision continue to have the same proportions as in the one repair action model. The choice between quick or slow repair depends critically on the relative outcomes of the two repair processes. However, the numerical example suggests one is more likely to move from slow to quick repairs as the environment becomes more hostile but this need not always be the case.

6.2 Stochastic Game approach

In chapter 4, We looked at conflict situations where the environment is controlled by an opponent. In this case the opponent's actions force the need for the equipment, and this situation is modelled as a stochastic game. For these stochastic game models, we developed models with global and local constraints on effort. These models have sought to investigate the maintenance and repair policy for a stand-by system where the environment of when it is needed is controlled by an opponent. The most obvious context for this problem is the military one either in conventional or peace keeping roles. In this research, we looked at the problem where there are no constraints on the enemy in terms of the actions they can perform. This leads to a complete but unrealistic solution to the problem because in reality, an opponent- be it rogue country or terrorist - cannot be continually on the attack. Thus we introduced the idea of a constraint on the average effort undertaken by the opponent over the total history of the game so far. We naively described this as a sleep index that the opponent needs to sleep for a certain percentage of the time. We also expanded these results to the situation where the advantage of a rest or quiescent period is discounted the further in the past it is, but always has a positive effect. In addition we looked at games where the benefit to the opponent of being " able to sleep" only lasts for a finite period and is then lost completely. In each case we derived properties of the form of the optimal maintenance policy which hold on all occasions and also found the form of the optimal policy in specific numerical examples.

One interesting feature is that the optimal policies may well be mixed so at each period there is a certain probability one should perform maintenance, and a certain probability one does nothing. Clearly if there are a number of such stand-by units, the mixed policy can translate into what proportion should be given preventive maintenance at that time. If the difference of initiating event probabilities between the benign and the dangerous environment (b_1, b_2) is small, one tends only to perform maintenance when system is close to failure, but in other situations one will maintain a very good system because one feels the environment is likely soon to be dangerous(high sleep index). As with the case in chapter 3, one always repairs a failed unit, no matter what the environment, since you can be sure that the enemy will seek to take advantage of the failure by increasing the danger in the environment if they can. The models introduced in chapter 4 are the first to address the question of maintenance in an environment where failure can be catastrophic and where there is an enemy seeking such catastrophes.

6.3 Training Model

Finally, as an extension of the models in chapter 3 we added training factor to the stand-by unit operator for deciding the optimal action and developed discrete time Markov decision process formulations for this problem. Equipment can only be used if it is in an operable state and if its users have had sufficient recent training with it. Thus as well as repairing and maintaining the equipment, it is necessary to train users. However, a problem with training is that it increases the wear and tear of the stand-by unit even though it enhances the operator's ability to respond well to an initiating event. We looked at the interaction between the need for training and the need to service the equipment. In order to do this, we developed a Markov decision process model without no loss of learning in the training level and a modified model where the effect of training does wear off. For training model with continuous loss of expertise, we developed an "expertise index" and looked at when one should train, when one should repair as a function of the environmental situation, the training level and the state of the equipment.

The results show that there is also a strong interaction between the quality state of the stand-by unit, the general environment state, the training level of the operator and the decision on whether to repair or train. In the training model with random loss of expertise, the expected survival time until a catastrophic situation decreases as the unit quality states, the training levels and environment situations worsen. One always repairs when the unit is down. Also, once repair is optimal at certain quality state, repair is always optimal for worse quality states. When the unit quality state is new, the training level is at its lowest which means that it can not respond to an initiating event and there is no difference between between deteriorations caused by natural factors and training, one always trains or repairs. One trains if the quality state is good and there is same quality level below which one repairs. If one adds the extra condition $\sum_{j<l} P_{1j} \geq \sum_{r<l} R_r$ to the above condition, one can show one always trains when the quality state is new if the training level is at its lowest. If the repair action is not available and the quality state is in working condition, once training is optimal at a certain training level, one always trains at worse training levels. In the training model with continuous loss of expertise, we find that one always repairs when the unit is down. If the unit is operating then as the expertise index increases, one moves from training to doing nothing, but this need not always be the case.

6.4 Future Development

We have studied repair strategies of stand-by units under changing environment in this thesis. For this we have used Markov decision process and stochastic game approaches. The models proposed in this thesis are a kind of prototype models. So we can expand and develop these prototype models to more complicated and sophisticated ones for future research. The followings are some suggestions for doing this.

Firstly we can consider more realistic quality state transition probabilities which vary according to the environment. In this thesis we have considered the fixed quality state transition probabilities for our Markov decision process and stochastic game models. This means that the variation of the quality state to the next decision epoch is not affected by the current environment situation. However, we could consider a generalisation where the transition matrix has different probabilities based on different environments.

Secondly we can think of more sophisticated dynamics of the changes in quality state and environment situation. For our models, we assumed that the variation of the quality state and environment situation are made only at fixed epochs $t = 0, 1, 2, \cdots$. However in many real repair and maintenance problems the times between consecutive variation epochs are not identical but random. Thus, we can consider semi-Markov decision models. These semi-Markov decision models describe dynamic systems which are observed and change at random points in time.

We could also expand our models to more realistic ones by giving more general repair and training times. In our models we have assumed repairing the equipment takes only 1 unit time period and 2 unit time periods for slow repair model. However if we consider more general times for the repair action, it is certain that we can have more realistic models. This can be applied to training as well. In all the above cases, it would be sensible to use real data from the field to check whether the assumptions and extensions suggested are valid in reality.

In the stochastic game models, we have developed the game models as 'Full

Game' models in which the players can know the other player's readiness perfectly. But in reality it is hard for one player to know the other's readiness one hundred percent. The players may have only partial information about each other. So we can develop more complicated game models which give only partial knowledge of the state of equipment to the players. We could also consider models where both players need to make repair decisions which affect the readiness.

Bibliography

Arrow, K. J., Blackwell, D., and Girshick M. A., 1949. Bayes and minimax solutions of sequential decision problems, Econometrica 17, 213-244

Arrow, K. J., Harris, T., and Marschack, J., 1952. Optimal inventory policy, Econometrica 19, 250-272

Barlow, R. E., Proschan, F., 1965. Mathematical theory of reliability, John Wiley, New York

Baker, R. D., Wang, W., 1992. Estimating the delay-time distribution of faults in repairable machinery from failure data. IMA J. Math. Appl. Bus. Ind., 3, 259-281

Baker, R. D., Wang, W., 1993. Developing and testing the delay time model. JORS, 44, 361-374

Baker, R. D., Christer, A. H., 1994. Review of delay-time OR modelling of engineering aspects of maintenance. European Journal of Operational Research, 73, 407-422

Bellman, R. E., 1957. Dynamic Programming, Princeton University Press, Princeton, USA

Bellman, R. E., Blackwell, D., 1949. On particular non-zero sum game, RM-250, Rand Corp., Santa Monica, USA

Bellman, R. E., LaSelle, J. P., 1949. On non-zero sum games and stochastic processes, RM-212, Rand Corp., Santa Monica, USA

Bewley, T., Kohlberg, E., 1976. The asymptotic theory of stochastic games. Math. Oper. Res. 1, 197-208

Binmore, K. 1992. Fun and game: A text on Game Theory. Houghton Mifflin

Blackwell, T., Ferguson, T.S., 1968. The Big Match., Ann. Math. Statist. 39, 159-163

Butler, D. A., 1979. A hazardous inspection model. Management Science 25, 79-89

Cao, J., 1988. Reliability analysis of a repairable system in a changing environment subject to a general alternating renewal process, Microelectronics and Reliability 28(6), 889-892.

Cao, J., 1989. Stochastic behaviour of a man-machine system operating under changing environment subject to a Markov process with two states, Microelectronics and Reliability 29(4), 529-531.

Cayley, A., 1875. Mathematical questions with their solutions, No. 4528, Educational Times 23, 18

Christer, A. H., 1982. Modelling inspection policies for building maintenance. JORS. 33, 723-732.

Christer, A. H., Redmond, D. F., 1990. A recent mathematical in maintenance theory. IMA J. Math. Appl. Bus. Ind., 2, 97-108.

Christer, A. H., Waller, W. M., 1984a. Delay time models of industrial inspection-maintenance problems, JORS, 35, 401-406.

Christer, A. H., Waller, W. M., 1984b.An operational research approach to planned maintenance: modelling PM for a vehicle fleet, JORS, 35, 967-984.

Christer, A. H., Wang, W., 1995. A delay time based maintenance model of a multi-component system, Math. Appl. Bus. Ind., 6, 205-222.

Christer, A. H., Wang, W., Choi, K., and Van Der Duyn Schouten, F. A., 2001. The robustness of the semi-Markov and delay time single-component inspection models to the Markov assumption, IMA Journal of Management Mathematics 12, 75-88.

Çinlar, E., 1984. Markov and semimarkov models of deterioration. In Reliability Theory and Models(M. Abdel-Hamid, E. Çinlar, J. Quinn, eds), Academic Press, New York, 3-41.

Çinlar, E., Özekici, S., 1987. Reliability of complex devices in random environments. Prob. Engrg. Inform. Sci. 3, 97-115.

Çinlar, E., Shaked, M. and Shanthikumar, J.G., 1989. On lifetimes influenced by a common environment. Stoch. Proc. Appl. 33, 347-359.

Dekker, R., 1996. Applications of maintenance optimization models: A review and analysis. Reliability Engineering and System Safety 51(3), 229-240.

Deng, Y., Song, S., 1993. Reliability analysis of a multicomponent system in a multistate Markovian environment. Microelectronics and Reliability 33(9), 1237-1239

Deng, Y., Tang, B., 1994. Comparing of system reliabilities of repairable coherent systems in a changing environment. Microelectronics and Reliability 34(2), 317-321

Derman C., 1963. On optimal replacement rules when changes of state are Markovian in Mathematical Optimization techniques , ed. by R. Bellman, Univ. of California Press, Berkely, 201-210.

Dhillon, B.S., 1982. Stochastic models for predicting human reliability. Microelectronics and Reliability 22(3), 491-496

Dhillon, B.S., Rayapati, S.N., 1985. Reliability evaluation of electric power systems in alternating environment. Microelectronics and Reliability 25(2), 357-367

Dhillon, B.S., Rayapati, S.N., 1987. Reliability evaluation of outdoor electric power systems. International Journal of Energy System 7(2), 85-89

Dhillon, B.S., Natesan, J., 1986. Stochastic analysis of two-unit outdoor electric power systems in changing weather. Microelectronics and Reliability 26(4), 679-690
Dvoretsky, A., Kiefer, J. and Wolfowitz, J., 1952. The inventory problem: I. Case of known distribution of demand, Econometrica 20, 187

Gessford, J., Karlin, S., 1958. Optimal policies for hydroelectric operations, in Studies in the mathematical Theory of Inventory and Production, edited by K. J. Arrow, S. Karlin, and H. Scarf, Stanford University Press, Stanford, USA

Gillette, D. 1957. Stochastic games with zero stop probabilities. Contributions to the Theory of Games, Vol. III(Annals of Mathematics Studies, No. 39). Princeton, N.J., 179-187

Goel, L.R., Gupta, R., Singh, S.K., 1985. Cost analysis of a two-unit standby system with delayed replacement and better utilization of units. Microelectronics and Reliability 25(1), 81-86

Goel, L.R., Kumar, A., Rastogi, A.K., 1985. Stochastic behaviour of man-machine systems operating under different weather conditions. Microelectronics and Reliability 25(1), 87-91

Goel, L.R., Gupta, R., Rastogi, A.K., 1985. Cost analysis of a system with partial failure mode and abnormal weather conditions. Microelectronics and Reliability 25(3), 461-466

Goel, L.R., Sharma, G.C., Gupta, R., 1985. Cost analysis of a system with partial failure mode and abnormal weather conditions. Microelectronics and Reliability 25(3), 461-466

Guo, T., Cao, J., 1992. Reliability analysis of a multistate one-unit repairable system operating under a changing environment, Microelectronics and Reliability 28(6), 889-892

Hastings, N. A. J., van Nunen, J.A.E.E., 1977. The action elimination algorithm for Markov decision processes. Ed. Tijms, H. C., Wessels, J. Mathematical Centre Tract, No. 93, Amsterdam, 161-170.

Heyman, D. H., Sobel, M. J.,1984. Stochastic Models in Operations Research, Vol. II. McGraw-Hill, New York, USA

Howard, R. 1960. Dynamic Programming and Markov Processes, MIT Press, Cambridge, USA

Isaacs, R., 1955. Games of Pursuit, P-257, Rand Corp., Santa Monica, USA

Karlin, S., 1955. The structure of dynamic programming models, Naval Res. Log. Quart. 2, 285-294

Kiefer, J., 1953. Sequential minimax search for a maximum, Proc. Amer. Math. Soc. 4, 502-506

Kiessler, P.C., Klutke, G.-A. and Yang, Y., 2002. Availability of periodically inspected systems subject to Markovian degradation. J. Appl. Prob. 39, 700-711.

Klutke, G.-A., Wortman, M. and Ayhan, H., 1996. The availability of inspected systems subject to random deterioration. Prob. Engrg. Inform. Sci. 10, 109-118.

Lefévre, C. and Milhaud, X. 1990. On the association of the lifelengths of components subjected to a stochastic environment. Adv. App. Prob. 22, 961-964.

Luss, M. and Kander, Z., 1974. Inspection policies when Duration of Checking is Non-negligible, Operational Research Quarterly 25, 299-309

Luss, M., 1976. Maintenance policies when deterioration can be observed by inspections, Operations Research 24, 359-366

McCall, J.J., 1965. Maintenance policies for stochastically failing equipment: A survey. Management Science 11(5), 493-524

Mertens, J. F., Neyman, A., 1981. Stochastic Game. Int. Journal of game Theory 10(2), 53-66.

Monahan, G.E., 1982. A survey of partially observable Markov decision processes: theory, models and algorithms. Management Science 28, 1-16.

Murthy, D. N. P., Yeung, V., 1995. Modelling and Analysis of Maintenance Service Contracts. Mathl. Comput. Modelling Vol. 22. No. 10-12, 219-225.

Murthy, D. N. P., Asgharizadeh, E., 1999. Optimal decision making in a maintenance service operation. European Journal of Operational Research 116, 259-273.

Özekici, S., 1995. Optimal maintenance policies in random environments. European Journal of Operational Research 82, 283-294.

Özekici, S., 1996. Complex systems in random environments. In Reliability and Maintenance of Complex Systems(S. Özekici, ed), Springer Verlag, New York, 137-157.

Pillay, A., Wang, J., Wall, A. D., and Ruxton, T., 2001. A maintenance study of fishing vessel equipment using delay-time analysis. J. Qual. Maint. Eng. 7, 118-127

Pierskalla, W.P., Voelker, J.A., 1976. A survey of maintenance models: the control and surveillance of deteriorating systems. Naval Research Logistics Quarterly 23, 353-388

Putterman, M. L., 1994. Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley & Sons

Rao, S. S., Chadrasekaran, R., and Nair, K. P., 1973. Algorithms for discounted games. J. Opt. Theory and Appl. 11, 627-637

Robins, H., 1952. Some aspects of the sequential design of experiments, Bull. Amer. Math. Soc. 58, 527-536

Rosenfield, D., 1976. Markovian deterioration with uncertain information. Operations Research 24, 141-155

Ross, S. M., 1983. Introduction to stochastic dynamic programming. Academic Press, New York

Sengupta, B., 1980. Maintenance policies under imperfect information. European Journal of Operational Research 5. 198-204.

Shaked, M. and Shanthikumar, J. G., 1989. Some replacement policies in a random environment.. Prob. Engrg. Inform. Sci. 3, 117-134

Shapley, L.S., 1953, `Stochastic Games', Proc. Nat. Ac. Sci. 39, 1095-1100

Sherif, Y.S., Smith, M.L., 1981. Optimal maintenance models for systems subject to failure-A review. Naval Research Logistics Quarterly 28(1), 47-74

Stalhane, T., 1990. Assessing software reliability in a changing environment. IFAC SAFECOMP'90, 83-88

Suzuki, K., On a preventive maintenance with monitored system. J. Oper. Res. Japan 22, 69-82

Thomas, L.C., Harley, R. and Lavercombe, A.C. 1983. Computational comparison of value-iteration algorithm for discounted Markov decision processes. Oper. Res. Lett. 2 (2), 72-76

Thomas, L.C., Harley, R. and Lavercombe, A.C. 1986. Computational comparison of policy iteration algorithm for discounted Markov decision processes. Comput. Oper. Res. 13(4), 411-420

Thomas, L.C., 1986. A survey of maintenance and replacement models for maintainability and reliability of multi-item systems. Reliability Engineering and System Safety 16 (4), 297-309

Thomas, L.C., 1986. Games, Theory and Applications. Ellis Horwood Limited

Thomas, L.C., Jacobs, P.A., Gaver, D.P.,1987. Optimal Inspection Policies for Standby Systems. Communication in Statistics-Stochastic Models 3(2), 259-273

Tijms, H. C., 1994. Stochastic Models: An Algorithmic Approach. John Wiley & Sons, New York.

Zhao-Tung, L., 1992. Reliability analysis of a repairable system in a randomly changing environment. Microelectronics and Reliability 32(10), 1373-1377

Valdez-Flores, C. Feldman, R.M., 1989. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. Naval Research Logistics 36, 419-446

van der Wal, J. 1977. Discounted markov games: successive approximations and stopping times. Int. J. Game Theory 6, 11-22

von Neumann, J., 1928, Zur theorie der gesellschaftsspield, Mathematische Annalen 100, 295-320

von Neumann, J., 1937, Über ein ökonomishes Gleichingssystem und eine Verallgemeinering des Brouwerschen Fixpunktsetzes, in: Ergebnisse eine Math. Coll. ed. K. Menger 8, 73-83

von Neumann, J. and Morgenstern, O. 1944, Theory of Games and Economic Behaviour(2nd edn, 1947), Princeton University Press, Princeton

Wald, A., 1947. Sequential Analysis, Wiley, New York, USA

Wang, R. C., 1977. Optimal replacement poicy with unobservable states. J. Appl. Prob. 14, 340-348

Wang, W., 1997. Subjective estimation of the delay time distribution in maintenance modelling. European Journal of Operational Research 51, 516-529

Wang, H., 2002. A survey of maintenance policies of deteriorating systems. European Journal of Operational Research 139, 469-489

Wattanapanom, N. and Shaw, L. 1979. Optimal inspection schedule for failure detection in a model where test hasten failures. Operations research 27, 303-317

Wartman, M. and Klutke, G.-A. 1994. On maintained systems operating in a random environment. J. Appl. Prob. 31, 589-594

White, C. C., 1979. Optimal control limit strategies for a partially observed replacement problem. Int. J. Systems. Sci. 10 , 321-331

Yang, Y. and Klutke, G.-A., 2000a. Improved inspection schemes for deteriorating equipment. Prob. Engrg. Inform. Sci. 14, 445-460

Yang, Y. and Klutke, G.-A., 2000b. Lifetime characteristics and inspection schemes for Lévy degradation processes. IEEE Trans. on Rel. 49, 337-382

Yeh, L. 1995. An optimal inspection-repair-replacement policy for standby systems. J. Appl. Prob. 32, 212-223.