63-2-2

Unclassified

AFESD - TDR - 62-265

Technical Report No. 279

R. C. Norris

Studies in Search for a Conscious Evader

14 September 1962

Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Unclassified

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with the joint support of the U.S. Army, Navy and Air Force under Air Force Contract AF 19(628)-500.

Non-Lincoln Recipients

PLEASE DO NOT RETURN

Permission is given to destroy this document when it is no longer needed.

Unclassified

263

AFESD - TDR - 62- 265

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINCOLN LABORATORY

STUDIES IN SEARCH FOR A CONSCIOUS EVADER

R. C. NORRIS

Group 22

TECHNICAL REPORT NO. 279

14 SEPTEMBER 1962

LEXINGTON MASSACHUSETTS

Unclassified

STUDIES IN SEARCH FOR A CONSCIOUS EVADER*

ABSTRACT

This ~~paper~~ analysis considers a search problem in which the search is directed against a conscious evader or an object controlled by a conscious evader. It is a two-person, zero-sum game called a search evasion game. Although the searcher cannot observe any of the evader's actions, the evader can observe the searcher's and can capitalize on errors that he makes.

At the beginning of the game, the evader hides in one of several boxes. The search process consists of a sequence of looks into the various boxes until the evader is found. Each look into a given box takes a fixed amount of time. If the searcher looks into the box in which the evader is located, he will find the evader with a certain probability – the detection probability associated with the box in question. A particular evasion device is assumed: the evader can move from one box to another between looks. A cost is usually associated with such a move.

Primary emphasis is placed on the study of the search evasion game that involves two boxes, for solutions have been found. Two limiting forms of the two-box game are considered first. in G^{∞}, moving is prohibited. In G°, the other limiting form, the evader can move at no cost.

The game becomes more interesting when a nonzero but finite cost is associated with each move. In most cases, a finite prohibitive bound on the moving cost exists. When the moving cost exceeds this bound, the searcher's good strategy is identical with his good strategy in G^{∞}. The evader should never move if the searcher uses this strategy. When the moving cost is strictly less than the prohibitive bound, the searcher's good strategy is Markovian in form. That is, the good search strategy can be generated by a finite Markov process in which a look is associated with each transition.

The search evasion game that involves more than two boxes is also studied. In G°, the limiting form in which the moving costs are equal to zero, exact solutions can still be found. The basic properties of the other limiting game, where moving is prohibited, are simple extensions of those that apply when there are only two boxes. In this game, however, the computational effort required to find a solution can be excessive.

The properties of the general many-box game in which the moving costs are neither prohibitive nor equal to zero are quite different from those that apply in the two-box case. Except when the moving costs are very small, the searcher's good strategy can no longer be generated by a Markov process. The complex character of the game is indicated by the partial solution that has been found to the simplest three-box game. The prospects of being able to find exact solutions to the general game in an efficient manner appear to be remote. A particular approach to finding approximately good search strategies is suggested for future research.

---

* This report is based on a thesis of the same title submitted to the Department of Electrical Engineering at the Massachusetts Institute of Technology on 31 August 1962, in partial fulfillment of the requirements for the degree of Doctor of Science.

ACKNOWLEDGMENT

I wish to express my deep appreciation to Professor R.A. Howard, the supervisor of the thesis on which this report is based. His suggestions and ideas were extremely helpful. Without his enthusiasm and encouragement, this work could never have been completed. I am also indebted to Professors P.M. Morse and E. Arthurs for their interest and useful comments.

TABLE OF CONTENTS

STUDIES IN SEARCH FOR A CONSCIOUS EVADER

CHAPTER 1
INTRODUCTION

1.1 HISTORY OF THE PROBLEM

Operations research was first recognized as a formal discipline during World War II when scientific methods of analysis were applied to operational military problems. One of the first endeavors of this new discipline in this country was the development of satisfactory methods for searching for enemy submarines. The search theory that evolved considered a homogeneous environment in which the hidden object (submarine) was located. From this theory developed search patterns that optimized the probability of detection when it was assumed that the hidden object was either stationary or moving in some prescribed manner independent of the search effort. Since then, this work has been developed further, notably by Koopman.[1-3]

Search problems in which the environment cannot be approximated by a homogeneous one have also been considered by assuming a discrete environment. In these problems, the search effort consists of a sequence of looks into various boxes in which the object may be hidden. In many cases, the time required to examine a given box is fixed or "quantized."

Gluss[4] considers a problem of this type in the process of developing sequences for testing the various subassemblies of a complex system in order to find a faulty component. In his work, he assumes that on each look the searcher either locates the object (faulty part) or gains no information. In other words, the failure to find the object in a given box does not decrease the probability that the object is there.

Pollock[5] treats a similar problem, particularly the case in which there are only two boxes. He assumes that when the correct box is examined, the object is found with probability q_i, where q_i is called the detection probability of the box in question. After an unsuccessful look, the probability that the object is in the box just examined is decreased according to Bayes' rule. Many of Pollock's results are found in Chapter 2. In particular, he originated the approach used in Sec. 2.3.

Another discrete search problem is considered by Blackman.[6,7] He studies a problem in which one or more objects appear as the search process goes along.

A feature common to all these problems is that the object never moves from one box to another.

Examples can be found in both continuous and discrete search problems, however, where the object need not be stationary or move in some arbitrary manner. Rather, the object may be an intelligent evader who attempts to outwit the searcher by moving so as to increase his chances of escape. The search problem in such a situation should be treated as a game where the actions of the evader are taken into account. Dubbins[8] points out this consideration in a discussion of the treatment of military problems associated with tactics, pursuit, evasion, search, and the like when he says, "For the most part these have been treated as 'one-sided' problems; given prescribed behavior for an opponent, one seeks to optimize the result of his own actions. With the development of the theory of zero-sum, two-person games, however, it has been natural to seek extensions to the more realistic 'two-sided' problems in which each of the two participants is free to choose his actions from a non-trivial class of possible strategies." This is a valid statement, and several of the problems mentioned above have been considered from a two-sided point of view. The author, however, has been unable to find any papers in which the theory of two-sided search is treated.

This paper will consider a discrete two-sided search problem in which the looks are quantized. To avoid the ambiguity of the term "two-sided search," it will be called a "search evasion game." Before going into the details of the game to be studied, let us consider a particular example.

1.2 THE REVENUER VS THE MOONSHINER: AN EXAMPLE

In a particular section of the hills of Tennessee, it is known that a moonshiner is operating an illegal still. As a result, a federal agent has been dispatched to the scene and a game ensues. In this game, we assume that the moonshiner can operate his still in any of several locations or areas known to both players. Each day, the revenuer selects one of these areas to search, and he continues his hunt until he catches his man. Since the moonshiner is a clever fellow, he can conceal his apparatus in such a manner that he will not necessarily be found when the revenuer searches the area in which he is located. Rather, he will be found with a certain probability, the detection probability of that area. It will be assumed that both players know the detection probability of each of the various areas.

The moonshiner, being a rational businessman at heart, is mainly interested in securing a good profit for himself. He knows that his still will yield him a profit of one unit per day. Through spies, or by other means, he can observe where the revenuer looks, and he realizes that he can prolong the expected length of his operation by changing the location of his still from time to time. Since the revenuer searches during the day, the evader knows that he can move his still with relative safety at night. However, when he moves it, he must suspend his operation and destroy the material being processed at the time. This will cost him μ units. We shall assume that once he has completed his move, he can immediately replace the process materials so that his future production is not affected. Nonetheless, our entrepreneur realizes that he suffers a loss in profit whenever he moves and that he must balance this moving cost against the advantage of a possibly longer career.

The revenuer is also an intelligent man who has specific motives. Since none of the areas has a detection probability equal to zero, he knows that he can eventually catch his man. However, in addition to catching criminals, he is interested in making this crime as unattractive as possible. That is, he considers deterrence to be one of his primary functions. As a result, he is interested in minimizing the expected profits which the moonshiner accumulates before he is caught. Therefore, the two players' interests are directly opposed.

This is a two-person, zero-sum game. The revenuer is a searcher and the moonshiner is an evader. Each of the areas in which the evader can hide can be called a box, and each box has an associated detection probability. The search process consists of a sequence of looks into the various boxes. The evader can observe where the searcher looks, and between each pair of looks the evader can move from one box to another. The game continues until the evader is found.

1.3 THE SEARCH EVASION GAME MODEL

The game described above is a particular example of the search evasion game to be studied in this paper. This game was motivated by a problem involving inspection under an arms control agreement. Assume that the manufacture of certain weapons systems is prohibited by an arms control treaty and that under this treaty an inspectorate is established to enforce the agreement. One of the functions of this inspectorate would be to visit the various places where such systems could be manufactured. The purpose would be twofold: to discourage a violation of the treaty and to disclose any such violation if it occurs. Although the inspectee may choose to honor the treaty, the inspector has no reason to assume this. Furthermore, whether the inspector wishes to deter a violation or minimize the possible advantages of such a violation, it is reasonable to assume that his opponent's gain is his loss. This results in a two-sided, zero-sum game in which the inspector should assume that a clandestine operation exists. When it does, the game becomes interesting.

Although the game to be studied was motivated by the arms control problem, no claims are made as to the validity of the model to be defined, in this context. Many simplifying assumptions have been made. Furthermore, all political considerations have been ignored and an arbitrary utility function is assumed. The result is a game that is studied for its own sake. It is a two-sided extension of a more classic one-sided search problem. It will be interesting to see how a particular evasion device – moving between looks – affects the behavior of the game. If the results of this study can be applied to a practical problem, perhaps the one just mentioned, so much the better.

In our search evasion game there are two players, the searcher and the evader. The evader must hide at the beginning of the game in one of a set of boxes. The searcher must make a sequence of looks into these boxes until he finds the evader. A look into a particular box takes a particular amount of time, known to both players. If the searcher looks into box i and the evader is there, the evader will be found with probability q_i, where q_i is the detection probability of box i. The detection probability of each box is known to both players. We shall always assume that the evader can observe where the searcher looks. Unless a statement is made to the contrary, we shall also assume that the evader can move from one box to another between looks. If a cost is associated with such a move, this cost is known to both players.

To complete the definition of this game, a utility function over the possible outcomes of the game is needed for each player. It is not appropriate to develop the theory of utility here. A good treatment of this theory as well as the theory of games in general can be found in Ref. 9. We shall assume that the utility of any outcome can be expressed in a numerical form equivalent to money. The game is zero-sum: the sum of the two players' utilities for a given outcome must equal zero. Thus, one player's utility is the negative of the other's, and only one of the utilities need be considered explicitly. In this game, the evader's utility will always be used.

The search evasion game is of a sequential nature and may be thought of as a two-sided extension of a sequential decision process. A particular play or outcome of the game consists of a particular sequence of events. An unsuccessful look into box i while the evader hides in box j is such an event. Similarly, an event occurs when the evader moves from one box to another or when he chooses not to move between a given pair of looks. Associated with each event is a "reward." This reward is equal to the amount that the event in question contributes to the evader's utility. Thus, the utility of a given play of the game is equal to the sum of the rewards associated with the events that occur. In the example of the revenuer vs the moonshiner, a reward of one unit was associated with each look. This reward did not depend upon the box in which the evader was hiding and was even collected on the final look when the evader was found. A reward was also associated with each move. This reward was equal to $-\mu$. If a reward is negative, we may refer to the corresponding positive quantity as a "cost" or a "loss." Thus, in the above example, a cost of μ was associated with each move.

Both players may use strategies involving random decisions. Also, a stochastic element is introduced by the detection probabilities of the various boxes. As a result, a particular play, or sequence of events, cannot be associated with a given pair of strategies for the two players. Rather, such a pair of strategies defines a probability distribution over the various plays that can occur. By taking the expected value of the utilities associated with these plays over the above probability distribution, a utility can be associated with a given pair of strategies. The evader's utility will be called the "payoff" for the given pair of strategies. This payoff will be equal to the expected value of the sum of the rewards associated with the various events that can occur. Since the payoff is equal to the total amount (in the expected sense) that the searcher pays to the evader, we see that the evader is interested in maximizing the payoff while the searcher is interested in minimizing it.

We shall assume that payments are made while the game is being played. When an event occurs, the searcher pays the appropriate reward to the evader. This will be necessary when discounting is considered in Chapter 7. The actual transfer of a reward, however, cannot be used to provide information concerning the event that occurred. Thus, the searcher cannot infer that a move occurred because he received μ units. We think of the rewards as being transferred from the searcher to the evader only because the searcher considers the evader's gain to be his loss.

In Chapters 2 through 7, the search evasion game that involves only two boxes will be studied in some detail, for the "good" strategies, or at least "ϵ-good" strategies, can be found for the two players.

In Chapters 2 through 5, a simple reward structure is assumed. The evader receives one unit for each look and pays μ units for each move. In Chapter 2, the game in which moving is prohibited (where the evader chooses only where he hides) is considered. In Chapter 3, the other limiting form of the game is considered – the game in which μ, the moving cost, is equal to zero. In Chapters 4 and 5, the more general game where μ is finite but unequal to zero is treated. In Chapter 4, the evader's good strategy is developed; in Chapter 5, the searcher's.

A more general reward structure is assumed for the two-box game in Chapter 6. The reward associated with a given look depends on where the look is made and also on where the evader is hiding. The moving cost depends on which move occurs. Finally, a detection loss is subtracted from the evader's payoff when he is found. This loss can depend on where the evader is hiding when detection occurs.

In Chapter 7, discounting is introduced. With discounting, the utility of a given event decays exponentially as the amount of time that elapses before the event occurs increases. Discounting is useful in situations where immediate rewards are more important than rewards delayed until the distant future.

In Chapter 8, the N-box form of the search evasion game is considered. When moving is prohibited, the general properties of this game are simple extensions of those of the two-box form. The computational effort required to find the good strategies, however, becomes far more difficult. When the moving costs are all equal to zero, the game is fairly simple, and exact solutions can be obtained. The good search strategy of this game will prove to be of special interest. The general N-box game in which the moving costs are unequal to zero but finite becomes very complex and the general approach used in the two-box game breaks down. The partial solution of a very simple example indicates the complex character of the general N-box game. It is believed that only approximate solutions are feasible, and a particular approach to finding an approximately good search strategy is suggested for future research.

CHAPTER 2

G^∞: THE SEARCH EVASION GAME WITH MOVING PROHIBITED

2.1 INTRODUCTION

In this chapter we shall study the search evasion game in which the evader is not allowed to move between looks. We should expect that in G, the game without this restriction, the evader would not move if the moving cost μ became sufficiently large. The game where moving is prohibited will, therefore, be called G^∞. As a limiting form, its properties should be of interest, and we shall see that the mathematical techniques developed in its study will be useful in the more general game.

The rules for G^∞ can be stated simply. The evader may hide in either of two boxes or may make any random choice between them. The searcher must pay the evader one unit for each look and must continue his search until he finds the evader. The payoff of the game for a given pair of strategies is equal to the expected search time. Associated with each box i is a detection probability q_i known to both players. It is the probability that a look into that box will reveal the evader's presence if he is there.

A strategy for the evader can be defined by the probability vector $\underline{P} = \{p_1, p_2\}$ that determines where he hides. The evader is not required to reveal the vector he uses to the searcher. A fundamental theorem from the theory of games states, however, that if both players use good strategies neither suffers any disadvantage by revealing his strategy to the other. As a result, we shall study the modified game F^∞ in which the searcher is informed of the vector that the evader selects. This study will compose the bulk of this chapter, for once the solution of F^∞ is known the solution of G^∞ is simple.

2.2 MODIFIED GAME F^∞: EVADER'S STRATEGY KNOWN TO SEARCHER

Having required the evader to reveal his strategy, F^∞ almost ceases to be a game. Once the evader has hidden, he no longer has any control over his fate, and the searcher is simply faced with a problem of optimization. With this as a rationale, we will adopt the convention of using the term "optimum strategy" in place of "good strategy" in reduced games of this type where one player is required to reveal his strategy to the other. The term "good strategy" will be reserved for use in the original game in which this requirement does not apply.

An optimum strategy for the searcher in F^∞ may be thought of as a rule for generating, as a function of $\underline{P}$, a search sequence that minimizes the expected search time. Since we are considering only two boxes, $\underline{P}$ has only one degree of freedom, and it is convenient to adopt the notation $\underline{P} = \{P, 1 - P\}$. The symbol P equals the probability that the evader will hide in box 1, and we can let $U^\infty(P)$ represent the expected search time that results when the searcher uses an optimum strategy. As we shall soon see, the optimum search strategy is simple in that $U^\infty(P)$ need not be known numerically in order to determine this strategy.

On the other hand, the evader's optimum strategy in F^∞ (his good strategy in G^∞) consists in selecting that P at which $U^\infty(P)$ is a maximum. For this reason, and because $U^\infty(P)$ is needed in order to determine the good search strategy in G^∞, it must be calculated. This calculation is fairly involved and will be considered in some detail. The techniques developed will be useful in the more general game where moving is allowed at a cost.

In the process of looking for the evader, the searcher must make a sequence of decisions as to where he should look, until the evader is found. Since the evader may take no countermeasures once he has hidden, these decisions can be made in advance and can be deterministic. Since the game ends with the first successful look, an optimum strategy may be viewed as associating with each P a single infinite look sequence that is used as long as necessary.

Although we can expect the optimum search sequences to be different for different values of P, it is worthwhile to see how the expected search time, or payoff, associated with a fixed sequence behaves as a function of P. Let a represent the payoff that results if the evader actually hides in box 1 (P = 1) and let b represent the corresponding payoff if the evader hides in box 2 (P = 0) when a fixed sequence is used. With this sequence, the payoff as a function of P is $[a\,P + b(1 - P)]$ and is linear. Consider the ensemble of linear functions generated by the infinite set of all infinite search sequences. The expected search time $U^{\infty}(P)$ must be the greatest lower bound on this ensemble. Therefore, $U^{\infty}(P)$ must be continuous and convex. Throughout this paper a function f(x) will be considered convex if for all $x_i, x_j \in X$, $0 \leqslant y \leqslant 1$,

$$yf(x_i) + (1 - y)\,f(x_j) \leqslant f[yx_i + (1 - y)\,x_j] \quad .$$

In many cases, we shall find that $U^{\infty}(P)$ is piecewise linear over any interval $(\epsilon, 1 - \epsilon)$ where $\epsilon > 0$. That is, if we exclude the intervals $(0, \epsilon)$ and $(1 - \epsilon, 1)$ from the interval (0, 1) over which P is defined, the remainder $(\epsilon, 1 - \epsilon)$ can be partitioned into a finite set of nonzero intervals over each of which $U^{\infty}(P)$ is linear. The quantity ϵ must be strictly greater than zero, for the linear intervals will always become arbitrarily short as P approaches zero or one. If $U^{\infty}(P)$ is linear over a nonzero interval, a single infinite search sequence is optimum over this interval, and $U^{\infty}(P)$ equals the payoff associated with this sequence over this interval. At a point where $U^{\infty}(P)$ is formed by the intersection of two linear functions, the sequences associated with both functions are optimum.

Some additional properties of $U^{\infty}(P)$ are easily shown. Since at least one look is required to find the evader, $U^{\infty}(P)$ must be positive. Furthermore, it must be bounded as long as the detection probabilities are all unequal to zero. This will always be assumed, since the game holds little interest otherwise. If the evader is known to be in a particular box, it is clear that the searcher should always look there. The expected search time is then equal to the reciprocal of the detection probability and we find that $U^{\infty}(0) = 1/q_2$ and $U^{\infty}(1) = 1/q_1$.

2.3 DYNAMIC PROGRAMMING WITH P AS A STATE VARIABLE

The function $U^{\infty}(P)$ has been defined as the payoff of F^{∞} that applies when the evader hides in box 1 with probability P and in box 2 with probability (1 − P) and the searcher uses an optimum search sequence. Since the searcher knows the a priori probability distribution defining the evader's position at the beginning of the game, he can calculate the a posteriori probability distribution that applies after a sequence of unsuccessful looks. Therefore, at any point in the game, we can use $\underline{P} = (P, 1 - P)$ to represent the probability distribution defining the evader's position at that time. The searcher's future behavior should depend only upon the value of P which applies at a given time. That is, the searcher's future sequence of looks should be the same as the entire sequence that would apply if the evader had originally used this P in hiding at the beginning of the game. The probability P can be treated as a state variable and $U^{\infty}(P)$ can be used to represent the future payoff that results if the searcher uses his optimum strategy for all future looks.

The manner in which P changes during the search process is easily shown. Let us adopt the notation $P \xrightarrow{i} P'$ to indicate that P is transformed into P' by an unsuccessful look into box i. Then,

$$P \xrightarrow{1} P' = \frac{Pr_1}{Pr_1 + (1 - P)} \quad , \qquad r_1 = 1 - q_1 \quad ; \tag{2-1}$$

$$P \xrightarrow{2} P' = \frac{P}{P + (1 - P)\, r_2} \quad , \qquad r_2 = 1 - q_2 \quad . \tag{2-2}$$

If a sequence involving more than one look transforms P into P', this sequence can be written over the arrow in a similar manner. However, the final transformation depends only on the total number of looks into each box and not on their order, and it is often convenient to represent the transformation that involves k_1 looks into box 1 and k_2 looks into box 2 by $P \xrightarrow{(k_1, k_2)} P'$. Then,

$$P \xrightarrow{(k_1, k_2)} P' = \frac{Pr_1^{k_1}}{Pr_1^{k_1} + (1 - P)\, r_2^{k_2}} \quad . \tag{2-3}$$

In the above expressions, r_1 and r_2, the complements of the detection probabilities q_1 and q_2 are used. We shall call these complements the escape probabilities. If the evader is hiding in box i and the searcher looks into that box, the evader will escape detection (will not be found) with probability r_i. Although only the detection probabilities are needed, we shall find it convenient to use both the r's and the q's in our expressions, with the condition $q_i + r_i = 1$ implied.

We are now in a position to write the fundamental functional equation. If the searcher looks into box 1, the evader receives one unit for the look and survives this look with probability $(Pr_1 + 1 - P)$. When this occurs, P transforms into $Pr_1/(Pr_1 + 1 - P)$. Similar conditions apply if the searcher looks into box 2. Letting $U^\infty(P; i)$ represent the payoff if box i is examined first, after which an optimum search strategy is employed, we have

$$U^\infty(P) = \min \begin{cases} U^\infty(P; 1) = 1 + [Pr_1 + 1 - P]\, U^\infty \left[\dfrac{Pr_1}{Pr_1 + 1 - P}\right] \\[2ex] U^\infty(P; 2) = 1 + [P + (1 - P)\, r_2]\, U^\infty \left[\dfrac{P}{P + (1 - P)\, r_2}\right] \end{cases} \quad . \tag{2-4}$$

This functional equation makes two problems apparent. In order to find the searcher's optimum strategy we must find which of $U^\infty(P; 1)$ and $U^\infty(P; 2)$ is smaller for a given P. Once this is known, we are still faced with the problem of evaluating $U^\infty(P)$, for Eq. (2-4) expresses $U^\infty(P)$ as a function of another unknown $U^\infty(P')$. Unless P eventually transforms back into itself after a finite number of optimum looks, $U^\infty(P)$ must be evaluated by means of an infinite series.

At present we are in a position to derive the searcher's optimum strategy.† It is clear that if P = 1, the searcher should look into box 1, and if P = 0 he should look into box 2. It is reasonable to assume that there exists a P_0 such that if P is greater than P_0 the searcher should look into box 1, whereas if it is less than P_0 he should look into box 2. When $P = P_0$ we should

† A rigorous proof is contained in Appendix A. A more general form of F^∞ is treated, and its reading should be deferred until after one has read Chapters 6, 7 and 8.

expect that the searcher can look into either box. A look into box 1, however, decreases P and requires that the next look be into box 2. Similarly, if the first look is into box 2, P will become greater than P_0 and the next look should be into box 1. Letting $U^{\infty}(P; 12)$ represent the payoff when the searcher looks into box 1 and then into box 2, after which an optimum strategy is employed, we have

$$U^{\infty}(P; 12) = 1 + [Pr_1 + 1 - P] + [Pr_1 + (1 - P)\, r_2]\, U^{\infty}\left[\frac{Pr_1}{Pr_1 + (1 - P)\, r_2}\right] \quad . \qquad (2\text{-}5)$$

In a similar manner,

$$U^{\infty}(P; 21) = 1 + [P + (1 - P)\, r_2] + [Pr_1 + (1 - P)\, r_2]\, U^{\infty}\left[\frac{Pr_1}{Pr_1 + (1 - P)\, r_2}\right] \quad . \qquad (2\text{-}6)$$

Note that both equations contain the same expression

$$[Pr_1 + (1 - P)\, r_2]\, U^{\infty}\left[\frac{Pr_1}{Pr_1 + (1 - P)\, r_2}\right]$$

for the expected future payoff after the first two looks. The equation

$$U^{\infty}(P_0) = U^{\infty}(P_0; 12) = U^{\infty}(P_0; 21)$$

cancels these terms and reveals that $P_0 q_1 = (1 - P_0)\, q_2$, or

$$P_0 = \frac{q_2}{q_1 + q_2} \quad . \qquad (2\text{-}7)$$

The searcher's optimum strategy therefore requires that

$$\text{if } P > P_0 = \frac{q_2}{q_1 + q_2} \quad \text{look into box 1} \quad ;$$

$$\text{if } P < P_0 = \frac{q_2}{q_1 + q_2} \quad \text{look into box 2} \quad ;$$

$$\text{if } P = P_0 = \frac{q_2}{q_1 + q_2} \quad \text{look into either box} \quad . \qquad (2\text{-}8)$$

Noting that $\underline{P} = \{P, 1 - P\} = \{p_1, p_2\}$, we see that the searcher's optimum strategy requires him to look into the box for which $p_i q_i$ is the larger. That is, the searcher should make the choice that maximizes the probability of finding the evader on the next look.

This strategy is, in a sense, a deterministic form of a behavioral strategy, and it is worthwhile here to contrast behavioral strategies with pure and mixed strategies. Let us consider a game tree. Each node of the tree corresponds to a move for one of the players. Each branch extending from the node represents one of the possible alternatives that the player can select on that move. The nodes of the tree are partitioned into a set of information sets. For all nodes in a given information set, the same information concerning the past play is available to the player whose move it is. All moves in a given information set must have the same alternatives.

A pure strategy may be thought of as a list that specifies the alternative which should be selected in each information set. A mixed strategy specifies a probability distribution over the set of pure strategies. Thus, when a player uses a mixed strategy, he selects a pure strategy at the beginning of the game by means of this probability distribution. Once this selection has been made, he selects his alternatives throughout play deterministically. In our game, an infinite search sequence is a pure strategy and a random selection of an infinite search sequence is a mixed strategy.

In a behavioral strategy, a player associates with each information set a probability distribution for selecting his next alternative. Thus, when a behavioral strategy is used, random decisions are employed throughout play and not just at the beginning. Behavioral strategies are completely general as long as the players have perfect recall, which they do in this game.

In succeeding chapters, we shall find that we can characterize our information sets by means of state variables. The strategies that we shall develop will be formulated in terms of decision rules which are functions of these state variables. They will, therefore, be behavioral strategies. In F^{∞}, the searcher's optimum decision rule is a deterministic function of P and therefore, in a rigorous sense, is a pure strategy. As our study develops, however, we shall find that behavioral strategies are employed increasingly.

2.4 EXAMPLE WHERE BOXES ARE IDENTICAL: $q_1 = q_2 = q$

When the two boxes are identical the searcher's strategy is very simple. $P_0 = 1/2$ and the searcher should always look where the evader is most likely to be. We also have the good fortune in this example to find that if the searcher looks first into one box and then into the other, P returns to its initial value, that is, $Pr_1/[Pr_1 + (1 - P)\, r_2] = P$. Clearly, if such a pair of looks is optimum, it should be repeated, and the total optimum sequence should consist of alternate looks into the two boxes.

In order to find when such a sequence is optimum, let us define P_{01} and P_{02} as the probabilities into which P_0 is transformed by an unsuccessful look into box 1 and box 2, respectively. Then,

$$P_0 \xrightarrow{1} P_{01} \equiv \frac{P_0 r_1}{P_0 r_1 + 1 - P}$$

and

$$P_0 \xrightarrow{2} P_{02} \equiv \frac{P_0}{P_0 + (1 - P)\, r_2} \quad .$$

These probabilities will be of use in the more general case where $q_1 \neq q_2$, and the definition given above should be kept in mind. In this example, however, $P_{01} = r/(1 + r)$, $P_{02} = 1/(1 + r)$, and we find that

$$P_{01} \xrightarrow{2} P_0 \quad ,$$

$$P_{02} \xrightarrow{1} P_0 \quad .$$

If P belongs to the interval (P_{01}, P_0) the searcher should look into box 2, and P transforms into (P_0, P_{02}). A look into box 1 is then called for, since $P > P_0$, and P transforms back to its original value. Therefore, the sequence 2121... is optimum when P belong to (P_{01}, P_0). Similarly, the sequence 1212... is optimum when P belongs to (P_0, P_{02}). Since

$$U^{\infty}\left[\frac{Pr_1}{Pr_1 + (1 - P)\, r_2}\right] = U^{\infty}(P) \quad ,$$

Eqs. (2-5) and (2-6) may be used to calculate the respective payoffs, which are

$$U^{\infty}(P) = \frac{2}{q} - (1 - P) \quad , \quad P \in (P_{01}, P_0) \quad ,$$

$$U^{\infty}(P) = \frac{2}{q} - P \quad , \quad P \in (P_0, P_{02}) \quad .$$

When P lies outside the interval (P_{01}, P_{02}), the optimum search sequence no longer consists of alternate looks into the two boxes. Rather, there must be a sequence of looks into one box until P enters this interval. At this point, an alternating sequence commences. Since P must eventually enter (P_{01}, P_{02}) where $U^{\infty}(P)$ is known, we can calculate the payoff outside this interval also.†

First, let us consider the case where P is greater than P_{02} and define P_k as the probability that transforms into P_0 after k looks have been made into box 1. Thus,

$$P_k \xrightarrow{(k,\, 0)} P_0$$

and

$$P_k = \frac{1}{1 + r^k} \quad .$$

Clearly, if P belongs to (P_k, P_{k+1}), it will be transformed into (P_0, P_{02}) by k looks into box 1, and

$$U^{\infty}(P) = 1 + (Pr + 1 - P) + (Pr^2 + 1 - P) + \ldots$$

$$+ (Pr^k + 1 - P)\, U^{\infty}\left(\frac{Pr^k}{Pr^k + 1 - P}\right) \quad ,$$

where

$$U^{\infty}\left(\frac{Pr^k}{Pr^k + 1 - P}\right) = \frac{2}{q} - \frac{Pr^k}{Pr^k + 1 - P} \quad .$$

Therefore,

$$U^{\infty}(P) = \frac{2}{q} + k - P\left(\frac{1 - r^{k+1}}{q} + k\right) \quad , \quad P \in \left(P_{k_1} P_{k+1}\right) \quad .$$

The optimum sequence calls for k looks into box 1 followed by 1212... .

When P is less than P_{01}, a sequence of looks must first be made into box 2. As a result of the symmetry of this problem, we can simply write

---

† If P = 0 or 1, P never changes and all looks are made into box 1 or box 2, respectively.

$$U^{\infty}(P) = \frac{2}{q} + k - (1 - P)\left(\frac{1 - r^{k+1}}{q} + k\right) \quad , \quad P \epsilon \left(P_{-k-1}, P_{-k}\right) \quad ,$$

where

$$P_{-k} \xrightarrow{(0,\,k)} P_0 \quad ,$$

or

$$P_{-k} = \frac{r^k}{1 + r^k} \quad .$$

The optimum sequence in the interval (P_{-k-1}, P_{-k}) consists of k looks into box 2 followed by 2121. . . .

This completes the solution. The function $U^{\infty}(P)$ is graphed in Fig. 1 for the case where $q = 1/2$.

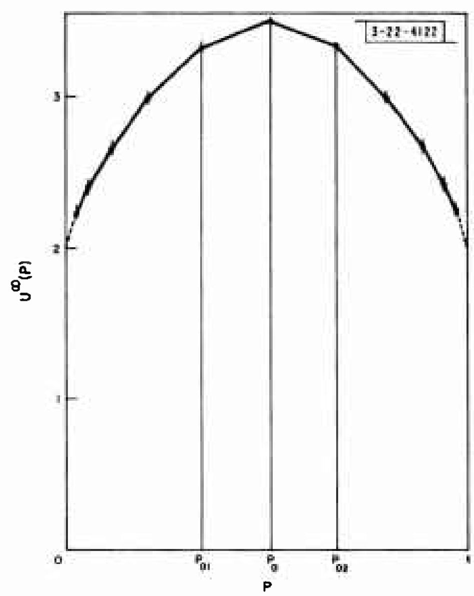


Fig. 1. $U^{\infty}(P)$: $q_1 = q_2 = 1/2$.

2.5 FURTHER PROPERTIES OF F^{∞}

Although the previous example is rather special, many of its characteristics are typical of the more general case where $q_1 \neq q_2$. In this more general case, $U^{\infty}(P)$ is convex. If two sequences are optimum at the same point P, this point must, in general, be transformed into P_0 by both sequences. Until this transformation occurs (for the first time), both sequences must be identical. In the case where $U^{\infty}(P)$ is piecewise linear over $(\epsilon, 1 - \epsilon)$, such a point must be a breakpoint where two linear intervals intersect. The optimum sequences associated with both intervals must be optimum at this point. In the previous example, P_3 was a breakpoint and it was transformed into P_0 by three optimum (and unsuccessful) looks into box 1. The optimum sequences associated with the linear intervals (P_2, P_3) and (P_3, P_4) were $\underline{111}2121\ldots$ and $\underline{111}1212\ldots$, respectively. If P is a breakpoint, it transforms into another breakpoint, whereas if P lies within a linear interval, the next optimum look transforms it into the interior of another linear interval. In general, P_0 will always be a breakpoint; that is, the derivative of $U^\infty(P)$ has a nonzero jump at this point. This occurs because the optimum next look is different on either side of P_0. The point P_0 is transformed into either P_{01} or P_{02}. We shall soon see that $U^\infty(P)$ is piecewise linear over any interval $(\epsilon, 1-\epsilon)$, where $\epsilon > 0$, if and only if P_{01} and P_{02} are transformed back into P_0 by finite sequences of optimum looks.

The behavior of P is interesting in that once P belongs to the interval (P_{01}, P_{02}) it remains in this interval for all time if optimum looks are used. If P is less than P_0, a look into box 2 is required and P increases to a new point that cannot exceed P_{02}. Similarly, if P is greater than P_0, a look into box 1 is required, and although P decreases in value it must still be larger than P_{01}. In contrast to this behavior, if P lies outside (P_{01}, P_{02}) it will eventually transform into it. If P is greater than P_{02} this transformation is accomplished by a sequence of looks into box 1. If P is less than P_{01}, a sequence of looks into box 2 serves the purpose (the bounding points $P = 0$ and $P = 1$ are exceptions because they transform into themselves). As a result of these properties, (P_{01}, P_{02}) will be called the recurrent region and the interiors of $(0, P_{01})$ and $(P_{02}, 1)$ will be called the transient regions.

The recurrent region is of special interest for several reasons. First, as we have seen from our example, once $U^\infty(P)$ is known within this region it is not too difficult to compute $U^\infty(P)$ for any P lying outside it. In addition, $U^\infty(P)$ attains its maximum somewhere inside this region. The point P^* at which this occurs corresponds to the evader's good strategy in G^∞. Also, we shall find that the infinite search sequences and the associated payoff functions optimum at P^* are all that are needed to derive the searcher's good strategy in G^∞.

The proof that $U^\infty(P)$ is a maximum inside (P_{01}, P_{02}) is quite simple. At P_0 a look into either box is optimum and we can write

$$U^\infty(P_0) = U^\infty(P_0; 1) = 1 + (P_0 r_1 + 1 - P_0)\, U^\infty(P_{01})$$

$$= U^\infty(P_0; 2) = 1 + [P_0 + (1 - P_0)\, r_2]\, U^\infty(P_{02}) \quad .$$

It follows that $U^\infty(P_{01}) = U^\infty(P_{02})$. The function $U^\infty(P)$ is convex, and furthermore, it can be shown that it cannot be flat over the whole recurrent region. Therefore P^* must lie inside the recurrent region.

2.6 PERIODIC BEHAVIOR IN THE RECURRENT REGION

In the previous example we found that there were two intervals inside (P_{01}, P_{02}) over which $U^\infty(P)$ was linear and that the optimum search sequence associated with each was periodic. This was true because, with the detection probabilities equal, a look into each of the two boxes transformed P back into itself and looks of this form were optimum in (P_{01}, P_{02}). With the periodicity that resulted, P could oscillate during the search process only between two points, one in each of the two intervals. Finally, since P transformed back into itself after two looks, it was easy to compute $U^\infty(P)$ in closed form for each of the two intervals.

In this section, we shall study the conditions under which P transforms into itself after n_1 looks have been made into box 1 and n_2 looks have been made into box 2. We shall find that when these conditions hold, some ordering of n_1 looks into box 1 and n_2 looks into box 2 will be optimum for each P belonging to the recurrent region. Furthermore, over this region $U^\infty(P)$ will consist of $n_1 + n_2$ linear segments, with n_1 in (P_0, P_{02}) and n_2 in (P_{01}, P_0). The point P will transform into each of these intervals over which $U^\infty(P)$ is linear before it returns to its starting point. We shall see that within each interval the associated optimum sequence is periodic with period $n_1 + n_2$ and that the periodic sequences for the various intervals differ from one another only in phase.

The condition under which P transforms into itself after a total of n_1 looks have been made into box 1 and n_2 looks have been made into box 2 is quite simple. Recalling Eq. (2-3),

$$P \xrightarrow{(n_1, n_2)} \frac{P r_1^{n_1}}{P r_1^{n_1} + (1 - P)\, r_2^{n_2}} \quad ,$$

we see that this transformation occurs if

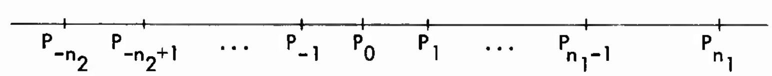
$$r_1^{n_1} = r_2^{n_2} \quad , \tag{2-9}$$

where n_1 and n_2 are integers, and r_1 and r_2 are the escape probabilities. Since we are interested in the first return, n_1 and n_2 should have no common factor. Equation (2-9) is equivalent to requiring that

$$\frac{n_1}{n_2} = \frac{\log(r_2)}{\log(r_1)} \quad .$$

A pair of integers (n_1, n_2) exists if the ratio of the logarithms of the escape probabilities is rational. If this ratio is irrational there still exist rational numbers that are arbitrarily close to it, and the case where $r_1^{n_1} = r_2^{n_2}$ is of general interest. The only exception occurs when r_1 and/or r_2 is equal to zero. This case is quite simple to solve and will be considered later. The problem of approximating $\log(r_2)/\log(r_1)$ by an n_2/n_1 for which $n_1 + n_2$ is not too large will also be deferred. Here we shall consider the behavior of F^∞ when the equation $r_1^{n_1} = r_2^{n_2}$ is satisfied exactly.

In order to see how the optimum search sequences behave in (P_{01}, P_{02}) under this condition, let us define an ordering of probabilities $P_{-n_2}, P_{-n_2+1}, \ldots P_{-1}, P_0, P_1, \ldots P_{n_1-1}, P_{n_1}$. Here, P_0 has its usual meaning. For the moment, we shall only require in addition that $P_{-n_2} = P_{01}$, $P_{n_1} = P_{02}$, and that $P_i < P_j$ if $i < j$:



If $P_i < P_j$, a look into box 2 will transform P_i to a point which is less than that into which P_j is transformed.

Making use of similar considerations, we can define the set $\{P_j\}$ by the following relations. Given P_i, a look into box 1 transforms it n_2 points to the left. Similarly, given a look into box 2, P shifts n_1 points to the right. Therefore, we can write

$$P_j \xrightarrow{(k_1, k_2)} P_{j-k_1 n_2 + k_2 n_1}$$

and we see that

$$P_j \xrightarrow{(n_1, n_2)} P_j \quad .$$

If P_j is greater than P_0, the searcher's strategy calls for a look into box 1, and if P_j is less than P_0 he should look into box 2. Let us consider a P_j that is unequal to P_{01}, P_0, or P_{02}. A little thought will show that since n_1 and n_2 have no common factor, P_j will eventually transform into P_0 as a result of an optimum sequence. At this point the searcher can look into either box and P can transform into P_{01} or P_{02}. In either case, the next look transforms P into $P_{n_1-n_2}$, and after a total of n_1 looks into box 1 and n_2 looks into box 2, P returns to P_j. Therefore, associated with each P_j are two optimum periodic sequences. These two sequences differ only in the order in which the two boxes are examined after P_j has transformed into P_0.

In order to calculate the actual value of an arbitrary P_j, we must note the number of looks into the two boxes (k_1, k_2) which transforms it into P_0. Then,

$$P_j \xrightarrow{(k_1, k_2)} P_0 = \frac{P_j r_1^{k_1}}{P_j r_1^{k_1} + (1 - P_j)\, r_2^{k_2}} \quad .$$

A simple manipulation reveals that

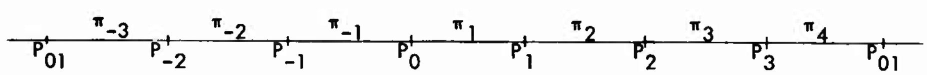
$$P_j = \frac{P_0 r_2^{k_2}}{P_0 r_2^{k_2} + (1 - P_0)\, r_1^{k_1}} \xrightarrow{(k_1, k_2)} P_0 \quad . \tag{2-10}$$

The entire set $\{P_j\}$ can be evaluated in this manner.

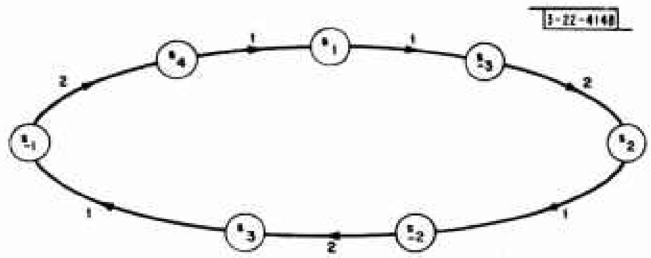
As has been mentioned, there are two optimum periodic sequences associated with each P_j. If we consider a P that lies just to the right of P_j we see that either sequence will transform P into the interval (P_0, P_1) as P_j transforms into P_0. A look into box 1 is required next, and the sequence associated with P_j that calls for this look will eventually transform P back into itself. Working through the same argument when P lies just to the left of P_{j+1}, we can conclude that the optimum periodic sequence common to P_j and P_{j+1} is the unique optimum sequence for all P inside (P_j, P_{j+1}). The function $U^{\infty}(P)$ consists of a single linear segment over the interval (P_j, P_{j+1}). The associated optimum sequence is periodic, with period $n_1 + n_2$, and one period of this sequence transforms P into each of the other intervals in (P_{01}, P_{02}) before transforming back into itself. Hence, the sequences associated with each interval are identical except in phase.

2.6.1 Example Where $r_1^4 = r_2^3$: Calculation of $U^{\infty}(P)$

Let us examine the case where $r_1^4 = r_2^3$. This will not only clarify the previous discussion, but will provide a device for showing how $U^{\infty}(P)$ can be calculated when the optimum search sequences are periodic inside the recurrent region. Ordering the breakpoints $P_{01} = P_{-3}$, P_{-2}, P_{-1}, P_0, $P_1, \ldots P_4 = P_{02}$ on the real line, we can designate each interval that results by π_i as indicated below.



In order to determine the periodic sequences associated with each interval it is convenient to draw a chain diagram showing the manner in which these intervals transform into each other. Each state s_i in the chain represents an interval π_i. Starting with π_1, a look into box 1 shifts P three (n_2) intervals to the left to π_{-3}, at which point a look into box 2 transforms it four (n_1) segments to the right to π_2. If the process is continued in this manner, the following diagram results. It is clear that one period of each optimum sequence involves n_1 looks into box 1 and



n_2 looks into box 2, and that these sequences differ only in phase. The sequence associated with s_2 is 1212112, 1212112, ... and so forth. In order to calculate the breakpoints it is convenient to note that $P_k \rightarrow P_0$ as $s_k \rightarrow s_{-1}$ when k is positive, and that $P_{-k} \rightarrow P_0$ as $s_{-k} \rightarrow s_1$. For example, $P_1 \xrightarrow{12121} P_0$, or $P_1 \xrightarrow{(3,\,2)} P_0$, so that

$$P_1 = \frac{P_0 r_2^2}{P_0 r_2^2 + (1 - P_0)\, r_1^3} \quad .$$

Once we have found the optimum sequence associated with an interval over which $U^{\infty}(P)$ is linear, we are in a position to calculate the payoff over this interval. This calculation is fairly straightforward. However, the general techniques that can be employed will be considered in some detail, since a thorough understanding of them is necessary when the more general game G, in which moving is allowed, is studied in Chapter 4.

The general approach that we shall use is to calculate the payoff which results when a particular search sequence is used. Associated with each state s_i in a chain diagram is a unique infinite sequence. We shall let $U_i^{\infty}(P)$ represent the payoff that results when this sequence is used. The payoff associated with s_i is, therefore, $U_i^{\infty}(P)$ and it is valid for all P belonging to the interval (0, 1). If the state s_i generates a sequence optimum over the interval π_i, then $U_i^{\infty}(P) = U^{\infty}(P)$ over this interval.

Usually, the state s_i will be assumed to generate a sequence that is optimum over π_i. Thus, we shall usually assume that $U_i^{\infty}(P) = U^{\infty}(P)$ when P belongs to π_i.

At times, however, the payoff of an approximately optimum search sequence will be considered. In this situation, $U_i^{\infty}(P)$ will be the exact payoff associated with s_i. The approximation that $U_i^{\infty}(P) = U^{\infty}(P)$ over π_i results from associating with this interval the approximately optimum search state s_i.

The payoff $U_i^{\infty}(P)$ is defined for a fixed sequence. This payoff, therefore, must be linear in P and we can express it in the form

$$U_i^{\infty}(P) = a_i P + b_i (1 - P) \quad .$$

Perhaps this is not the most obvious form in which the function could be expressed, but it will prove to be the most convenient. Although the evader may have chosen to hide in box 1 with probability P, once he has made this choice he knows exactly where he is. The quantity a_i is his expected payoff if he is actually in box 1, and b_i is his expected payoff if he is actually in box 2 when the searcher uses the sequence of s_i.

When this formulation is used, the manner in which the payoff of one state is related to that of the state to which it is connected by the next look becomes quite simple. If $s_i \xrightarrow{1} s_j$, we may use Eq. (2-4) to write

$$U_i^\infty(P) = a_iP + b_i(1 - P) = 1 + (Pr_1 + 1 - P)\, U_j^\infty\left(\frac{Pr_1}{Pr_1 + 1 - P}\right) \quad ,$$

which shows that

$$s_i \xrightarrow{1} s_j \Longrightarrow$$

$$a_i = 1 + r_1 a_j \quad ,$$

$$b_i = 1 + b_j \quad . \tag{2-11}$$

Similarly, we find that

$$s_i \xrightarrow{2} s_j \Longrightarrow$$

$$a_i = 1 + a_j \quad ,$$

$$b_i = 1 + r_2 b_j \quad . \tag{2-12}$$

If we wish to express $U_i^\infty(P)$ in terms of a $U_j^\infty(P)$ that follows after several looks, we may simply compound the above equations. However, it will prove to be more convenient to use an alternative formulation. In this formulation we can consider a complete set of mutually exclusive events, multiply the reward associated with each by the probability of the event, and sum. For example, if $s_i \xrightarrow{1211} s_j$ and the evader is in box 1, he can be found on the first, third, or fourth look or he can survive all of them. The rewards associated with these events are 1, 3, 4 and $4 + a_j$, respectively, and the associated probabilities of these events are q_1, $q_1 r_1$, $q_1 r_1^2$ and r_1^3. We can, therefore, write

$$a_i = q_1(1 + 3r_1 + 4r_1^2) + r_1^3(4 + a_j) \quad .$$

In order to write general expressions of this form, consider the set $\{t_m(n)\}$, where $t_m(n)$ represents the look on which box m is examined for the n^{th} time. In the above case, $t_1(1) = 1$, $t_1(2) = 3$, $t_1(3) = 4$ and $t_2(1) = 2$. If s_i is transformed into s_j by a sequence defined by $\{t_m(n)\}$ that involves a total of k_1 looks into box 1 and k_2 looks into box 2, then

$$s_i \xrightarrow{\{t_m(n)\}} s_j \Longrightarrow$$

$$a_i = q_1 \sum_{n=1}^{k_1} t_1(n)\, r_1^{n-1} + r_1^{k_1}(k_1 + k_2 + a_j) \quad ,$$

$$b_i = q_2 \sum_{n=1}^{k_2} t_2(n)\, r_2^{n-1} + r_2^{k_2}(k_1 + k_2 + b_j) \quad . \tag{2-13}$$

If $\{t_m(n)\}$ is used to represent the sequence that transforms s_i into itself, $k_1 = n_1$ and $k_2 = n_2$, and we find that

$$s_i \xrightarrow{\{t_m(n)\}} s_j \Longrightarrow$$

$$a_i = \frac{1}{1 - r_1^{n_1}} \left[q_1 \sum_{n=1}^{n_1} t_1(n)\, r_1^{n-1} + r_1^{n_1}(n_1 + n_2) \right] \quad ,$$

$$b_i = \frac{1}{1 - r_2^{n_2}} \left[q_2 \sum_{n=1}^{n_2} t_2(n)\, r_2^{n-1} + r_2^{n_2}(n_1 + n_2) \right] \quad . \tag{2-14}$$

As an illustration let us consider the interval π_1 in our example where $r_1^4 = r_2^3$. Referring to state s_1 in the chain diagram we find that one period of the associated optimum sequence is 1212121. Therefore,

$$a_1 = \frac{1}{1 - r_1^4} [q_1(1 + 3r_1 + 5r_1^2 + 7r_1^3) + 7r_1^4]$$

and

$$b_1 = \frac{1}{1 - r_2^3} [q_2(2 + 4r_2 + 6r_2^2) + 7r_2^3]$$

Once $U_1^\infty(P)$ is known, the payoffs for the other states can be computed in order around the chain by means of Eqs. (2-11) and (2-12). If only a particular payoff is desired, Eq. (2-13) may be used.

In Fig. 2 the payoff over the recurrent region is graphed for the case where $r_1 = 0.512$ and $r_2 = 0.4096$, i.e., where $r_1^4 = r_2^3 = x^{12}$ and $x = 0.8$.

These equations can also be used to compute the payoffs in the transient regions. However, the linear intervals approach zero in length as P approaches zero or one, and one should not attempt to calculate every payoff.

Equation (4-13) can be used to calculate the payoff for an arbitrary sequence even if it is not periodic. In this case, the set $\{t_m(n)\}$ must be used to represent the entire infinite sequence and

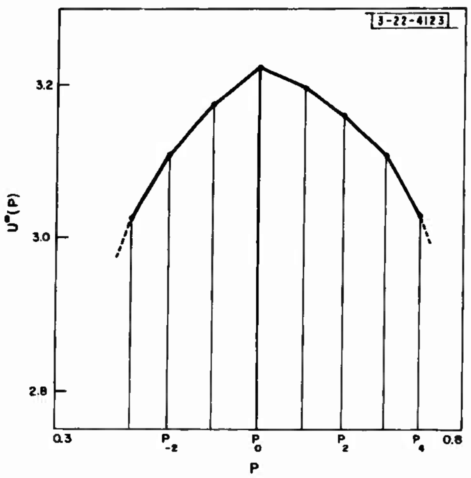


Fig. 2. $U^{\infty}(P)$: $r_1 = 0.512$, $r_2 = 0.4096$.

$$a = q_1 \sum_{n=1}^{\infty} t_1(n)\, r_1^{n-1} \quad ,$$

$$b = q_2 \sum_{n=1}^{\infty} t_2(n)\, r_2^{n-1} \quad .$$

If the payoff is known for any fixed sequence as a function of P, Eq. (2-13) may be used to compute the payoff that results if this sequence is preceded by an arbitrary finite sequence.

2.6.2 Applicability of the Periodic Case

It should be clear that for any pair of escape probabilities (r_1, r_2) where both are unequal to one, a pair of integers (n_1, n_2) can be chosen which makes n_1/n_2 arbitrarily close to $\log r_2/\log r_1$. The choice of n_1 and n_2 depends upon the accuracy desired in the approximation. The choice that is made will determine the sequences used and the payoff that results.

Even when $n_1 + n_2$ is fairly small, the approximation can be fairly good if n_1 and n_2 are well chosen. Associated with each search state s_i is an interval π_i. The breakpoints that define the boundaries of each such interval are those points that transform into P_0 sometime during the first $n_1 + n_2 - 1$ looks of the actual optimum sequence (P_{-n_2} and P_{n_1} should be calculated in this manner even though they will be only approximately equal to P_{01} and P_{02}, respectively). As a result, the first period of the sequence associated with s_i will be optimum for any P belonging to π_i. During the next few periods, errors can be made only when P is close to P_0, and their effect will be small. With increasing time, the errors become more frequent and more serious. The effect of these larger errors is mitigated, however, by the decreasing probability that the game lasts long enough for them to be made.

Perhaps the most intriguing aspect of an approximation of this type concerns the appearance of $U^{\infty}(P)$. Given a choice (n_1, n_2) the resulting payoff in $(P_{-n_2}, P_{n_1}) \cong (P_{01}, P_{02})$ will consist of $n_1 + n_2$ linear segments. Since the associated sequences are only approximately optimum, we cannot expect the function to be exactly continuous at the breakpoints. The interesting consideration, however, concerns the piecewise linear appearance of this function compared with that of the optimum payoff function. If $\log r_2/\log r_1$ is irrational, or if it is equal to n_1'/n_2' where $n_1' + n_2' \gg n_1 + n_2$, the optimum payoff function will have many more linear segments and dividing breakpoints. A breakpoint occurs at a point that transforms into P_0, and the sequences associated with the interval on either side will agree up to the time when this occurs. Breakpoints that transform quickly into P_0 will be much more apparent than those that transform into P_0 later on. If the integers (n_1, n_2) are well chosen, the important breakpoints will appear in the payoff function of the approximating strategy.

2.7 F^∞ WHEN q_1 AND/OR q_2 IS EQUAL TO ONE

When the detection probability of at least one box is equal to one, the optimum search strategy can never consist of a periodic sequence with looks into both boxes. Once a box with unity detection probability has been examined, it should never be examined again. If $q_1 = q_2 = 1$, the search can last for at most two looks, whereas if $q = 1$ for only one box, an optimum sequence can call for at most one look into it. In either case, the game is fairly easy to solve.

When both detection probabilities are equal to one, the game is trivial. If P is greater than one-half, the optimum search strategy calls for a look into box 1, followed by a look into box 2 if necessary. When P is less than one-half, the reverse is true. Clearly, $U^\infty(P) = P + 2(1 - P)$ for $P \geqslant 1/2$, and $U^\infty(P) = 2P + (1 - P)$ for $P \leqslant 1/2$.

When $q = 1$ for only one box, the optimum search strategy is almost as simple. Let us consider the case where $q_2 = 1$ and $q_1 \neq 1$. If P is less than P_0, box 2 should be examined first. P then becomes equal to one and all the future looks should be made into box 1. Since $U^\infty(1) = 1/q_1$, we find that $U^\infty(P) = P[1 + (1/q_1)] + (1 - P)$ when $P \leqslant P_0 = 1/(1 + q_1)$, and the payoff function consists of only one linear segment over $(0, P_0)$.

Over the interval $(P_0, 1)$, the payoff function consists of many segments. In fact, there are an infinite number of them, since as P approaches one they become arbitrarily short. For all P in a given interval over which $U^\infty(P)$ is linear, the same number of looks into box 1 is required before a look into box 2 is called for. We can define each breakpoint P_k by $P_k \xrightarrow{(k,\,0)} P_0$, and designate π_k as the interval (P_{k-1}, P_k). If P belongs to π_k, then k looks into box 1 transforms it to the left of P_0. Using Eq. (4-13) we find, after a simple manipulation, that

$$U_k^\infty(P) = P\left(\frac{1}{q_1} + r_1^k\right) + (1 - P)(k + 1) \quad ,$$

where

$$\pi_k = \left(\frac{1}{1 + q_1 r_1^{k-1}} \quad , \quad \frac{1}{1 + q_1 r_1^k}\right) \quad .$$

The function that applies when $q_1 = 1/2$ is graphed in Fig. 3.

It is worthwhile to note that there are dangers inherent in assuming that a detection probability is equal to one. Usually, if a search strategy that is optimum for one pair of detection probabilities is used for a slightly different pair, the payoff will be almost optimum. However, if one of the q's is assumed to equal one when it is only close to one this is no longer so. In this case, the searcher will look only once into that box. In the unlikely but possible event that the evader is in that box and escapes detection, he will receive an infinite payoff. This is clearly

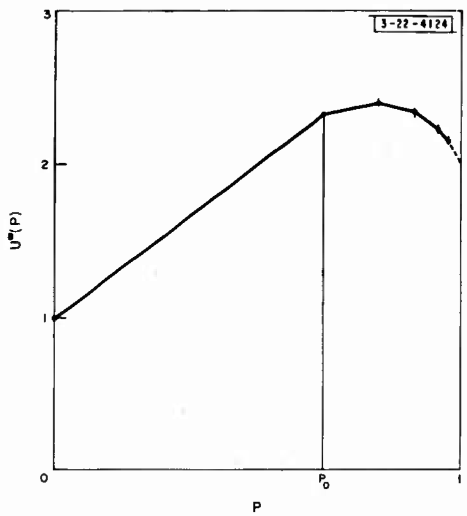


Fig. 3. $U^\infty(P)$: $q_1 = 1/2$; $q_2 = 1$.

undesirable. If both boxes are assumed to have unity detection probabilities the danger is not quite so great. If the evader is not found after the first two looks, the fallacy is revealed.

It should be mentioned, on the other hand, that $U^\infty(P)$, the payoff that applies when the actual optimum search strategy is used, is well behaved as one or both of the q's approach one. Therefore, if one wishes to get a rough idea of what the payoff is when q_i is close to one, the $q_i = 1$ solution is valid as an approximation.

2.8 SOLUTION OF G^∞

Now that we have considered the modified game, we are in a position to find the good strategies and value of G^∞ where the evader is not required to reveal his strategy to the searcher. In F^∞, a great deal of emphasis was placed on deriving the searcher's optimum strategy and the resulting payoff as a function of P. The evader's optimum strategy was scarcely mentioned because it was so simple – he should select the P at which $U^\infty(P)$ is a maximum, i.e., P^* where $U^\infty(P^*) \equiv V^\infty$. We now need to find the strategy for the searcher that limits the evader to this amount when P is unknown. This will be the searcher's good strategy, and it will imply that P^* is the evader's and that V^∞ is the value.

The searcher's good strategy is easily derived. We can usually expect that $U^\infty(P)$ will be a maximum at a unique point and therefore that the segments on either side will have slopes of opposite sign. Let the associated intervals be designated by π_i and π_j. If the searcher chooses to use the sequence associated with s_i with probability y_i and that associated with s_j with probability y_j, where $y_i + y_j = 1$, the payoff will be

$$U^\infty(P; y_i, y_j) = y_i U_i^\infty(P) + y_j U_j^\infty(P) \quad .$$

P can take any value between zero and one and is unknown to the searcher. Both $U_i^\infty(P)$ and $U_j^\infty(P)$ are equal to V^∞ at P^* and have slopes (with respect to P) of opposite sign. It follows that there exists a probability distribution (y_i, y_j) that yields a payoff equal to V^∞ for all P.

In the event that $U^\infty(P)$ is a maximum over a whole interval π_i where it has zero slope, the strategies are even simpler. The evader can select any P belonging to this interval, and the searcher's good strategy consists simply of the sequence that is optimum over π_i.

In order to illustrate the manner in which the searcher's good strategy is derived, let us consider the previous example in which $r_1 = 0.512$, $r_2 = 0.4096$, and where $r_1^4 = r_2^3$. The payoff $U^\infty(P)$, shown in Fig. 2, is a maximum at $P^* = P_0 = 0.5475$. (Although P^* is often equal to P_0, this is not always true.) The intervals on either side of P_0 are π_{-1} and π_1, which have payoffs

$$U_{-1}^\infty(P) = 3.549P + 2.827(1 - P) \quad ,$$

$$U_1^\infty(P) = 3.025P + 3.461(1 - P) \quad .$$

In general,

$$U^\infty(P; y_i, y_j) = (a_i y_i + a_j y_j)\, P + (b_i y_i + b_j y_j)\, (1 - P)$$

and is independent of P (and equal to V^∞) if $a_i y_i + a_j y_j = b_i y_i + b_j y_j$. Therefore, y_i must satisfy the equation

$$y_i = 1 - y_j = \frac{b_j - a_j}{(a_i - b_i) + (b_j - a_j)} \quad . \tag{2-15}$$

In our example we find that

$$y_{-1} = 0.377 \quad ,$$

$$y_1 = 0.623 \quad ,$$

and $V^\infty = 3.222$. The searcher's good strategy requires that he select the periodic sequence 2112121, 2112121... with probability 0.377 and 1212121, 1212121, . . . with probability 0.623. The evader should hide in box 1 with probability 0.5475 and in box 2 with probability 0.4525.

CHAPTER 3
G°: SEARCH EVASION GAME WITH ZERO MOVING COST

3.1 INTRODUCTION

In this chapter let us consider the other limiting form of our search evasion game – the game G° in which the evader can change boxes after each look with zero cost. Since μ, the moving cost, is equal to zero, the payoff from the searcher to the evader is again simply equal to the expected number of looks required to find the evader. Here, however, the evader plays an active role throughout the game. We should expect that, because of this additional freedom, he can guarantee himself a payoff larger than V^{∞}.

If we assume that the evader is playing against an intelligent searcher, we should expect him to make his moves in a judicious manner. The decision rule for making these moves should be formulated so as to accomplish a specific purpose – to maximize the guaranteed payoff. The evader should not move according to whim but only according to a carefully defined rule. Since the characteristics of such a decision rule will be common to the more general game where moves must be paid for, let us examine them in some detail.

3.2 A PROPERTY OF EVASION STRATEGIES; EVADER'S GOOD STRATEGY IN G°

In general, a moving strategy, or behavioral strategy, for the evader should specify as a function of past play a probability distribution for moving before the next look. Thus, at some point in the play, the evader may decide to move to box 1 if in box 2 with probability x_1 and to move to box 2 if in box 1 with probability x_2. In order to see what the effect of such a rule is, let us assume that we as observers know the past search sequence and all such decision rules that have been used up to this point. Given this information, we can calculate P, the probability that the evader is in box 1. By exercising the above decision rule, the evader transforms P into a new value, P'. Clearly,

$$P' = P(1 - x_2) + (1 - P)\, x_1$$

and

$$1 - P' = Px_2 + (1 - P)\,(1 - x_1) \quad ,$$

where only one of these equations is necessary. A result of such a strategy is, therefore, the transformation of the state variable P into P'.

In fact, this transformation is the only effect that the moving strategy has on the future behavior of the game. In order to show this, let us assume that the evader is required to reveal each decision rule of the above form to the searcher when it is exercised (note that we are not requiring him to reveal his complete strategy all at once). Such a revelation will not hurt the evader if he is using his good strategy and the searcher is intelligent enough to use his. On the other hand, the evader will certainly suffer a disadvantage if he uses a poor strategy, since the searcher can use the value of P' to determine where he should look next. This is the only manner in which the searcher can make use of this knowledge, and it follows that the only purpose of a moving strategy is this transformation.

We noted earlier that a moving strategy (x_1, x_2) could be a function of the entire previous play. However, given x_1 and x_2, the a posteriori P' is a function of the a priori P alone and not of the entire past play in all its detail. Therefore, the evader's good strategy must belong to the class of behavioral strategies in which the move probabilities x_1 and x_2 are functions of the a priori P. These arguments apply equally well to the general game where $\mu \neq 0$.

A strategy belonging to this class is completely defined by the functions $x_1(P)$ and $x_2(P)$ and the initial P that the evader uses when he first hides. The influence of the functions $x_1(P)$ and $x_2(P)$ on the behavior of the game may be described completely by the mapping of P that they produce. We shall determine the evader's good strategy in terms of this mapping. We can expect the mapping associated with the good strategy to be unique. For a given mapping, however, the functions $x_1(P)$ and $x_2(P)$ are not unique but have one degree of freedom. Therefore, the evader's good strategy will not be unique so far as these functions are concerned. In the next chapter, we shall see that when $\mu \neq 0$, the cost associated with a transformation of P depends upon the particular functions used. The functions that produce a given mapping at minimum cost are unique.

The evader's good strategy in G° is easily derived. In a manner analogous to that used in the last chapter, let us consider the modified game F° in which the evader must tell the searcher the value of P that applies before each look. In this game, the evader is completely free to change P after each look and therefore needs to consider only the effect that P has on the next look. If a given P is optimum before one look it should be used before each look. Therefore, after each look the evader should return P to its original position if it is optimum and we may write

$$U^\circ(P) = \min \begin{cases} U^\circ(P;1) = 1 + [Pr_1 + 1 - P]\, U^\circ(P) \\ U^\circ(P;2) = 1 + [P + (1-P)\, r_2]\, U^\circ(P) \end{cases} .$$

If a given look is optimum once, it is always optimum. Therefore,

$$U^\circ(P) = \min \begin{cases} U^\circ(P;1) = \dfrac{1}{Pq_1} \\ U^\circ(P;2) = \dfrac{1}{(1-P)\, q_2} \end{cases} .$$

The optimum P is that which maximizes $U^\circ(P)$ and is $P_0 = q_2/(q_1 + q_2)$. Thus, the evader hides in a manner that causes the probability of detection on each look to be independent of where the look is made. Using this optimum strategy, the evader guarantees a payoff in F° of $(1/q_1) + (1/q_2)$.

The functions $x_1(P)$ and $x_2(P)$ that achieve the optimum mapping are defined by the equation

$$P_0 = \frac{q_2}{q_1 + q_2} = P[1 - x_2(P)] + (1-P)\, x_1(P) ,$$

where we must require that $0 \leqslant x_1,\ x_2 \leqslant 1$. As long as the evader uses a good strategy throughout play, his strategy may be more simply defined. Given a look into box 1,

$$P_0 = \frac{q_2}{q_1 + q_2} \xrightarrow{1} P = \frac{q_2 - q_1 q_2}{q_1 + q_2 - q_1 q_2} .$$

Similarly,

$$P_0 = \frac{q_2}{q_1 + q_2} \xrightarrow{2} P = \frac{q_2}{q_1 + q_2 - q_1 q_2} .$$

Therefore, after each look into box 1, the evader should select x_1 and x_2 which satisfy

$$\frac{q_2}{q_1 + q_2} = \frac{(q_2 - q_1 q_2)(1 - x_2) + q_1 x_1}{q_1 + q_2 - q_1 q_2} ,$$

and after each look into box 2 he should select those which satisfy

$$\frac{q_2}{q_1 + q_2} = \frac{q_2(1 - x_2) + (q_1 - q_1 q_2) x_1}{q_1 + q_2 - q_1 q_2} .$$

Again, x_1 equals the probability of moving to box 1 if the evader is in box 2 and x_2 is the probability of moving to box 2 if he is in box 1.

In the next section we shall show that the searcher can limit the evader to the payoff $(1/q_1) + (1/q_2)$. This will prove that the above optimum strategy in F° is the evader's good strategy in G° and that $V^\circ = (1/q_1) + (1/q_2)$.

3.3 SEARCHER'S GOOD STRATEGY

The searcher's good strategy is also easily derived. In G^∞ the searcher's good strategy required each look to be a function of the past search sequence. He made a random choice between two infinite search sequences and after this choice was made all looks were specified deterministically. In G°, we find the opposite extreme. The evader can change P at will, and the searcher cannot make any use of his past sequence in choosing his next look. Therefore, we may consider only the class of behavioral search strategies in which each look is made independently of the others; that is, where box 1 is examined with probability Y and box 2 is examined with probability $1 - Y$.

It is convenient to turn the tables on the searcher and define the modified game H° in which he must reveal his probability distribution to the evader. Letting $W^\circ(Y)$ equal the payoff that results when the evader uses an optimum strategy, we find that

$$W^\circ(Y) = \max \begin{cases} W^\circ(Y; 1) = 1 + [Y r_1 + (1 - Y)]\, W^\circ(Y) \\ W^\circ(Y; 2) = 1 + [Y + (1 - Y)\, r_2]\, W^\circ(Y) \end{cases} ,$$

where $W^\circ(Y; i)$ is the payoff if the evader hides in box i. Clearly, if box i is optimum once it is always optimum. Therefore,

$$W^\circ(Y) = \max \begin{cases} W^\circ(Y; 1) = \dfrac{1}{Y q_1} \\ W^\circ(Y; 2) = \dfrac{1}{(1 - Y)\, q_2} \end{cases} .$$

The searcher's optimum strategy minimizes $W^\circ(Y)$, and we find that he should look into box 1 with probability $Y = q_2/(q_1 + q_2)$ and into box 2 with probability $1 - Y = q_1/(q_1 + q_2)$.

This strategy causes the probability of detection on any look to be independent of where the evader hides. It limits the evader to the payoff $(1/q_1) + (1/q_2)$, and is, therefore, the searcher's good strategy in G°.

It is interesting to note that in this game the payoff is equal to V° if only one player uses his good strategy. In G^{∞}, the payoff is always equal to V^{∞} if just the searcher uses his good strategy. If the evader uses his good strategy but the searcher uses a poor strategy, however, the payoff can be greater than the value.

CHAPTER 4
SEARCH EVASION GAME WITH $\mu \neq 0$: EVADER'S STRATEGY

4.1 INTRODUCTION

In the last two chapters we have considered the two limiting forms, G° and G^∞, of the two-box search evasion game. In G° the evader was completely free to move, and we found that after each look he should return the state variable P to the value that minimized the probability of detection on the next look. The game degenerated to a sequence of move-look pairs, each of which was independent of the previous ones (except that the game stopped once the evader was found). In G^∞ the opposite occurred. Once the evader hid, P became a function of the search process only. Therefore, the evader had to consider the influence of the whole search process on his choice of P. The searcher's good strategy became deterministic once an initial random selection was made from two infinite sequences. G^∞ was considered a limiting form of G because it appeared plausible to assume that the evader would never choose to move if the moving cost became infinite. We shall find that this is indeed true. In fact, we shall usually find that the evader should never move (even if the searcher is aware of this) as long as μ is larger than some finite μ_p. That is, when μ is greater than μ_p we may consider the moving cost prohibitive since the increase in search time achieved by moving is more than offset by the cost of moving.

In game G, when the moving cost is neither prohibitive nor zero we shall find characteristics intermediate between those of G° and G^∞. The evader receives one unit each time the searcher examines a box but must pay μ units each time he moves. Therefore, he must balance the increase in search time afforded by moving against the cost of moving. Also, we shall find that the searcher can make use of his past search sequence in determining where to look next but must make some random decisions throughout play. The good strategies for the two players, as would be expected, change in a well behaved manner from those associated with G° to those associated with G^∞ as μ increases. Furthermore, the value of the game decreases monotonically as μ increases from zero to μ_p. Once μ is greater than μ_p, the value is independent of μ and equal to V^∞ because the evader never incurs a moving charge.

In this chapter we shall develop the evader's good strategy. This will be accomplished by using the device that we used in the previous chapters – the modified game F in which the evader must reveal part of his strategy to the searcher. Many of the properties and techniques developed in studying F^∞ and F° will be useful here. As before, we shall proceed on faith and assume that the evader's optimum strategy in the modified game will be his good strategy in G. This faith will be justified, for we shall find in the next chapter that the searcher can indeed limit the evader in G to a payoff equal to that which the evader can guarantee himself in F. Also, as in the no-move game, we shall find that the optimum search strategies and the associated payoff functions developed in F will be of use when the searcher's good strategy in G is considered.

4.2 SOME RESTRICTIONS ON THE EVADER'S GOOD STRATEGY: EFFICIENT MOVE CONDITION

In the last chapter, we saw that the influence of the evader's strategy on the behavior of the game could be completely characterized by the manner in which it transformed the state variable P between looks. It followed that the evader's good strategy must belong to a class of behavioral strategies in which the probability of moving is a function of P and the box in which the evader finds himself. As a result, we were able to characterize the evader's good strategy in terms of a mapping and an initial P used at the start of the game.

When μ is unequal to zero, these properties still hold. As before, once the evader has exercised a moving strategy, the future payoff can be characterized by the a posteriori value of the state variable and the strategies that the players use in the future. Now, however, a cost is associated with a transformation of the state variable, and it is clear that if a given transformation of P into P' is desired, it should be achieved at minimum cost. This condition causes the move probabilities $x_1(P)$ and $x_2(P)$ associated with a given mapping to be unique and allows us to associate with any transformation a unique cost function $C(P \to P')$.

The move probabilities that achieve a given transformation at minimum cost are easily derived. If the desired transformation is $P \to P'$, then x_1 and x_2 must satisfy the equation

$$P' = P(1 - x_2) + (1 - P)\, x_1 \quad .$$

The probability that a move occurs is equal to $Px_2 + (1 - P)\, x_1$ and the cost of the transformation is simply $\mu[Px_2 + (1 - P)\, x_1]$. The quantities x_1 and x_2 must minimize this cost, subject to the usual restriction that $0 \leqslant x_1,\ x_2 \leqslant 1$, and we find that

$$P' < P \Longrightarrow x_1 = 0 \quad , \quad x_2 = \frac{P - P'}{P} \quad ;$$

$$P' > P \Longrightarrow x_1 = \frac{P' - P}{1 - P} \quad , \quad x_2 = 0 \quad . \tag{4-1}$$

This implies that the evader should never move from a box unless he wishes to decrease the probability that he is there. The cost of the transformation, which is now a minimum, is equal to $\mu(P - P')$ if $P' < P$ and $\mu(P' - P)$ if $P' > P$. We may therefore write

$$C(P \to P') = \mu\,|P - P'| \quad . \tag{4-2}$$

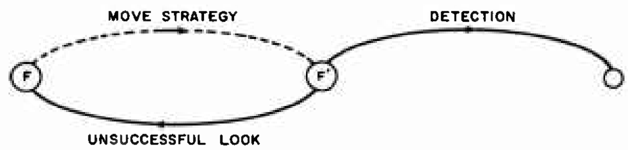
With the evader limited to strategies belonging to the class of efficient-move behavioral strategies defined above, we are in a position to develop the functional equations from which the good strategy can be computed.

4.3 MODIFIED GAMES F AND F': ASSOCIATED PAYOFF FUNCTIONS

In the last chapter, the evader's good strategy was derived by considering the modified game F° in which the evader was required to inform the searcher of the value of the state variable that applied before each look. We can do the same thing here. Now, however, a cost $C(P \to P')$ is associated with any transformation of the state variable and it will prove convenient to split the game into two parts. This will allow us to develop two payoff functions, one that applies before the evader exercises his move strategy and another that applies immediately thereafter.

These payoff functions can be developed by defining the modified games F and F'. In both games, the evader is required to reveal the value of P that applies before each look. Game F applies prior to the time when the evader exercises his moving strategy. Game F' applies after this has occurred. Therefore, we can consider F' as the game in which the evader is not allowed to move until after the next look has been made. Clearly, F and F' are two parts of a single sequential game. Their relation to each other is shown in the following diagram.

3-22-4149



In this diagram, the transition from F to F' that occurs when the evader exercises a move strategy is represented by a broken line, whereas the transition that occurs when the searcher makes an unsuccessful look is represented by a solid line. This convention will be used in all future chain diagrams, etc. Usually, however, any transition representing detection will be left out. Then, as in the chain diagrams used in Chapter 2, all look transitions will represent unsuccessful looks.

Payoff functions can be associated with both F and F' in the same manner as before. We can let U(P) represent the future payoff in game F when the evader is in box 1 with probability P and both players use optimum strategies in the future. Similarly, we can define U'(P) as the corresponding payoff in game F'. Note that P is used to represent the state variable in either game.

The functional equations that express these payoffs in terms of each other are easily derived. Given game F', the searcher must decide which box should be examined, and his decision can be based on the present value of P. The situation is exactly the same as in game F^∞ except that if the look is unsuccessful, game F is played next. Therefore,

$$U'(P) = \min \begin{cases} U'(P;1) = 1 + (Pr_1 + 1 - P)\, U\left[\dfrac{Pr_1}{Pr_1 + 1 - P}\right] \\ U'(P;2) = 1 + [P + (1-P)\, r_2]\, U\left[\dfrac{P}{P + (1-P)\, r_2}\right] \end{cases} \quad . \qquad (4\text{-}3)$$

As before, U'(P; i) represents the payoff that results if box i is examined and both players use optimum strategies thereafter.

In game F, the evader has the opportunity to transform P into some other P'. He must weigh the cost of such a transformation against the future payoff U'(P') in the subsequent game F'. Clearly,

$$U(P) = \max_{P'} \{-\mu\,|P' - P| + U'(P')\} \quad . \qquad (4\text{-}4)$$

These two functional equations are necessary and sufficient. That is, a unique pair of functions U(P) and U'(P) exists that satisfies the above equations. Once they are known, the optimum strategies for the two players can be found easily. This fundamental property, plus others of interest, is developed in Appendix B. Since most of the properties are fairly clear once they are stated, the proofs are not included here.

It is interesting to note in passing, however, that the proofs are accomplished by the use of the truncated games F_n and F'_n. Given F_n, the evader exercises a move strategy and F'_n is played. If the next look is unsuccessful, F_{n-1} follows and this process continues down to F_0. If the evader survives until this point, the game stops and he collects an additional reward that is independent of P. It is shown that as n approaches infinity, $U_n(P)$ and $U_n'(P)$ approach limiting forms U(P) and U'(P), respectively. Furthermore, it is shown that as n approaches infinity, the probability that the game will last until F_0 approaches zero. It follows that these limiting functions must satisfy Eqs. (4-3) and (4-4). Furthermore, most properties that hold for all F_n or all F_n' must apply equally well in F or F', respectively.

An important property developed in this manner is that both U(P) and U'(P) are continuous and convex. We shall see in Sec. 4.6 that they are also piecewise linear over the entire interval (0, 1) if the moving cost is not prohibitive. The quantity U'(P) is represented by a function of this form in Fig. 4.

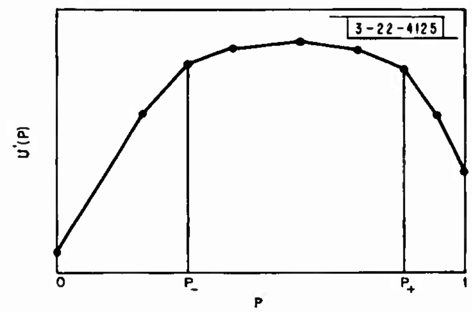


Fig. 4. A typical payoff function.

With Fig. 4 in mind, we are in a position to find the simple manner in which U(P) is related to U'(P) and the general form of the evader's optimum strategy. The points P_- and P_+ included in this figure are defined as follows:

$$
\begin{aligned}
\frac{dU'(P)}{dP} &\geqslant \mu \quad , \quad P \leqslant P_- \\
&< \mu \quad , \quad P > P_- \quad ; \\
-\frac{dU'(P)}{dP} &< \mu \quad , \quad P < P_+ \\
&\geqslant \mu \quad , \quad P \geqslant P_+ \quad .
\end{aligned} \tag{4-5}
$$

Equation (4-4) states that

$$U(P) = \max_{P'} \{-\mu|P - P'| + U'(P')\} \quad .$$

A little thought will show that if $P_- \leqslant P \leqslant P_+$, then U(P) = U'(P). On the other hand, if $P < P_-$, it follows that $U(P) = -\mu(P_- - P) + U'(P_-)$, and if $P > P_+$, then $U(P) = -\mu(P - P_+) + U'(P_+)$. If P lies in the interval (P_-, P_+), the evader should not move before the next look, and we shall call this interval the no-move region. If P is less than P_-, the evader should transform P to P_- by moving to box 1 if he is in box 2 with probability $x_1 = (P_- - P)/(1 - P)$. If P is greater than P_+, he should transform P into P_+ by moving to box 2 if he is in box 1 with probability $x_2 = (P - P_+)/P$. The intervals $(0, P_-)$ and $(P_+, 1)$ will be called moving regions.

In the next section we shall see that the magnitude of the slope of U'(P) must be greater than μ at P = 0 if $q_2 \neq 1$ and at P = 1 if $q_1 \neq 1$. Therefore, except in these unusual cases, P_- and P_+ belong to the interior of the interval (0, 1). These points are defined with care because U'(P) is usually piecewise linear, and it is possible for one of these segments to have a slope whose magnitude is exactly equal to μ.

The function U(P) can be constructed from U'(P) by replacing this function in the moving regions $(0, P_-)$ and $(P_+, 1)$ by tangent segments with slopes μ and $-\mu$, respectively. This is shown in Fig. 5. Clearly, U(P) achieves its maximum inside the no-move region.

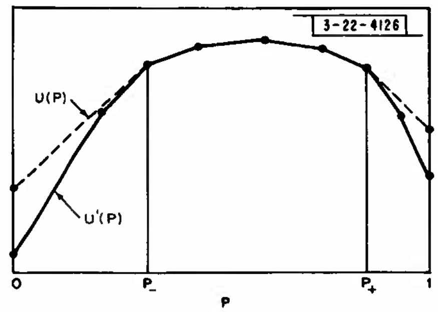


Fig. 5. The relationship between U(P) and U'(P).

The final property that is developed in Appendix B by means of the truncated games concerns the searcher's optimum strategy in F', namely:

There exists a P_0, where $P_- \leqslant P_0 \leqslant P_+$, such that

$$P < P_0 \Longrightarrow U'(P; 2) < U'(P; 1) \Longrightarrow \text{look into box 2} \quad ,$$

$$P > P_0 \Longrightarrow U'(P; 1) < U'(P; 2) \Longrightarrow \text{look into box 1} \quad .$$

Thus, the searcher's optimum strategy is quite similar to that in F^∞. Here, however, P_0 will in general be a function of the moving cost in addition to the detection probabilities. In F^∞, P_0 was simply equal to $q_2/(q_1 + q_2)$.

4.4 PROHIBITIVE MOVING COST

It has been mentioned that a finite bound μ_p usually exists above which the moving cost is prohibitive. That is, for any μ greater than μ_p, game G behaves in a manner essentially identical to that of G^∞. In particular, the value and the searcher's good strategy are identical to those in G^∞; the evader's good strategy requires him to hide initially in box 1 with the same probability P^* as in G^∞, and finally the evader should never move as long as the searcher uses his good strategy.

In this section, the conditions that insure this behavior will be considered, and we shall find that μ_p may be obtained from the payoff function $U^\infty(P)$. Furthermore, we shall find that the evader's complete good strategy, including the rule for moving when the searcher does not use a good strategy, can be easily obtained once $U^\infty(P)$ is known. It follows that one can determine whether μ is prohibitive and compute the good strategies from the solution of the no-move game.

Let us first consider more closely the conditions under which the evader should move. In the last section we found that the evader should exercise a moving strategy if P does not belong to the no-move region (P_-, P_+). Such a strategy transforms P to the nearest boundary of (P_-, P_+), and since the game should start at P^*, which lies inside this region, moving is required only when the state variable is transformed out of the no-move region by an unsuccessful look.

If P cannot be removed from the no-move region by an optimum look, the evader will never need to move when the searcher uses an optimum strategy.

In order to determine when this occurs, let us define the recurrent region (P_{01}, P_{02}) in the same manner as in F^{∞}. That is, let

$$P_0 \xrightarrow{1} P_{01} \equiv \frac{P_0 r_1}{P_0 r_1 + 1 - P_0} \quad ,$$

$$P_0 \xrightarrow{2} P_{02} \equiv \frac{P_0}{P_0 + (1 - P_0)\, r_2} \quad . \tag{4-6}$$

Since the searcher should look into box 1 if P is greater than P_0 and into box 2 if P is less than P_0, P can never be removed from (P_{01}, P_{02}) by an optimum look. Also, P can never be removed from this region by the evader if he uses his optimum strategy, because such an optimum strategy can only shift P in the direction of P_0 but not beyond it. Therefore, for any μ, the recurrent region has the same property that it had in F^{∞}.

The condition under which P cannot be removed from the no-move region by an optimum look should now be clear. The no-move region (P_-, P_+) must contain the recurrent region (P_{01}, P_{02}). That is, we require that $P_- \leqslant P_{01}$ and $P_{02} \geqslant P_+$. If P belongs to (P_{01}, P_{02}), it will remain there and hence inside $(P_- P_+)$. On the other hand, if P belongs to the no-move region but not to the recurrent region, it will soon be transformed into the recurrent region by means of a sequence of optimum looks. During this process, P moves towards (P_{01}, P_{02}) and therefore cannot leave the no-move region.

When this condition holds, U(P) must be equal to U'(P) throughout the recurrent as well as the no-move region. We may therefore derive P_0, the unique point at which a look into either box is optimum, in the same manner as it was derived in Chapter 2. Since

$$U'(P_0; 1) = U(P_0; 1) \quad ,$$

$$U'(P_0; 12) = U(P_0; 12) \quad , \qquad \text{etc.,}$$

the derivations are identical and

$$P_0 = \frac{q_2}{q_1 + q_2} \quad .$$

It follows that P_{01} and P_{02} have the same values as in F^{∞}. Furthermore, both U(P) and U'(P) are identical to $U^{\infty}(P)$ throughout the no-move region. In Sec. 4.3, the no-move region was defined as the interval over which $|[dU'(P)]/dP|$ is less than μ. Therefore, the no-move region contains the recurrent region and μ is prohibitive if and only if $|[dU'(P)]/dP|$ is less than μ for all P belonging to (P_{01}, P_{02}). Since $U^{\infty}(P)$ is convex, we may define μ_P by

$$\mu_p = \max \begin{cases} \left.\dfrac{dU^{\infty}(P)}{dP}\right|_{P_{01}+} \\[2ex] -\left.\dfrac{dU^{\infty}(P)}{dP}\right|_{P_{02}-} \end{cases} \quad . \tag{4-7}$$

The quantity μ_p will be finite as long as the magnitude of the slope of $U^\infty(P)$ is finite over the recurrent region. This will be true if both q_1 and q_2 are unequal to one, since both P_{01} and P_{02} belong to the interior of (0, 1) under this condition. When both detection probabilities equal one, $P_{01} = 0$ and $P_{02} = 1$. In this case, however, $U^\infty(P)$ consists of only two linear segments, both of finite slope, and $\mu_p = 1$. The only case where μ_p is infinite occurs when one, but not both, of the detection probabilities is equal to unity. For example, when $q_2 = 1$ and $q_1 \neq 1$, $P_{02} = 1$ and the magnitude of the slope of $U^\infty(P)$ approaches infinity as P approaches one.

A heuristic argument supports the fact that μ_p is infinite when just one detection probability is equal to one. If the searcher assumed that the evader could not move when μ was finite, he would look only once into box 2 ($q_2 = 1$). After this look, the evader would be willing to pay any finite price to move to box 2 if he were in box 1, since he could then survive for all time and collect an infinite payoff.

Since U(P) is equal to $U^\infty(P)$ over the entire no-move region when $\mu > \mu_p$, the evader's complete good strategy can be easily derived. The quantity U(P) is a maximum inside (P_-, P_+) and the state variable P^*, which should be chosen at the beginning of the game, and the maximum guaranteed payoff $U(P^*)$ are the same as in F^∞. Since P can still be transformed outside of the no-move region if the searcher does not use a good strategy, one must calculate the values of the points P_- and P_+. This can be done by finding the points at which the magnitude of the slope of $U^\infty(P)$ first exceeds μ. Note that both U(P) and U'(P) are unequal to $U^\infty(P)$ over the moving regions.

The searcher's good strategy in G is identical to that in G^∞ as long as μ is greater than μ_p. As we saw in Chapter 2, the good strategy consists of a random selection of the two infinite sequences that are optimum in F^∞ at P^*. Each of these sequences has an associated payoff function $U_i^\infty(P) = a_iP + b_i(1 - P)$. After each look, the future sequence is the same as that associated with another linear segment of $U^\infty(P)$ in (P_{01}, P_{02}) and can be represented in the same manner. For all of these sequences,

$$|a_i - b_i| = \left| \frac{dU_i^\infty(P)}{dP} \right|$$

must be less than μ. Therefore, even if the evader knows the future sequence, he will find it unwise to move. The initial random selection of one of the two optimum sequences associated with P^* is such that it limits the evader to a payoff independent of P and equal to $U^\infty(P^*) = U(P^*) = V^\infty$. When the searcher uses this good strategy the evader will receive this payoff as long as he does not move.

4.5 BEHAVIOR OF THE PAYOFF FUNCTIONS WHEN μ IS LESS THAN μ_p

In Sec. 4.3, some general properties of the payoff functions were discussed, and in the last section we saw that both of them were identical to $U^\infty(P)$ in the no-move region when the moving cost was prohibitive. In this section we shall consider the case where the moving cost is not prohibitive and examine the properties of the payoff functions more closely. Particular emphasis will be placed on the manner in which these functions change as μ increases from zero to μ_p.

Before considering the case where $0 < \mu \leqslant \mu_p$, however, it is worthwhile to consider the appearance of the payoff functions when μ is equal to zero. In Chapter 3 we found that the evader should always return the state variable to $P_0 = q_2/(q_1 + q_2)$ before each look and that he could guarantee a payoff equal to $(1/q_1) + (1/q_2)$. It follows that P_- and P_+ are both equal to P_0 and that the no-move region is simply a point. Since $\mu = 0$, the quantity $U(P)$ is simply equal to $U'(P_0)$ for all P and is thus a constant. In game F', however, the evader is not allowed to move before the next look. Therefore, the payoff $U'(P)$ will be a function of P. Using Eq. (4-3) and noting that $U(P) = (1/q_1) + (1/q_2) \equiv V^\circ$ for all P, we find that

$$U'(P) = \begin{cases} P(1 + r_1 V^\circ) + (1 - P)(1 + V^\circ) \quad , & P \geqslant P_0 \\ P(1 + V^\circ) + (1 - P)(1 + r_2 V^\circ) \quad , & P \leqslant P_0 \quad . \end{cases}$$

Both $U(P)$ and $U'(P)$ are shown in Fig. 6, where the convention of representing $U(P)$ by a broken line outside the no-move region is used.

If μ is very small but unequal to zero, the appearance of these functions can be only slightly different. Since μ is unequal to zero, $U(P)$ can no longer be flat but must decrease at a rate equal to μ as it extends from each side of the no-move region. For any fixed P, $U(P)$ and $U'(P)$ must be continuous functions of μ. They must be identical over the no-move region, and $U'(P)$ must be linear over an interval that is transformed into one of the moving regions by an optimum look. It follows that the no-move region must still consist of the point P_0 (which can change with μ) when μ is slightly greater than zero.

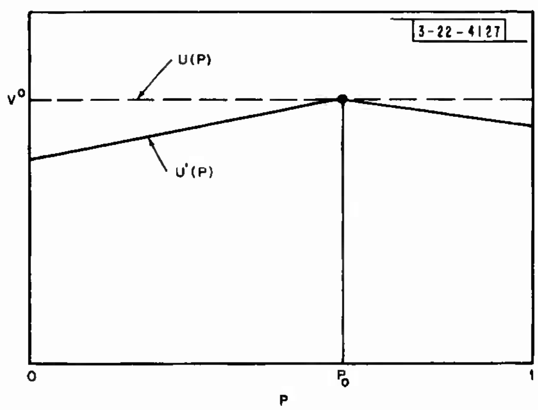


Fig. 6. Payoffs when $\mu = 0$.

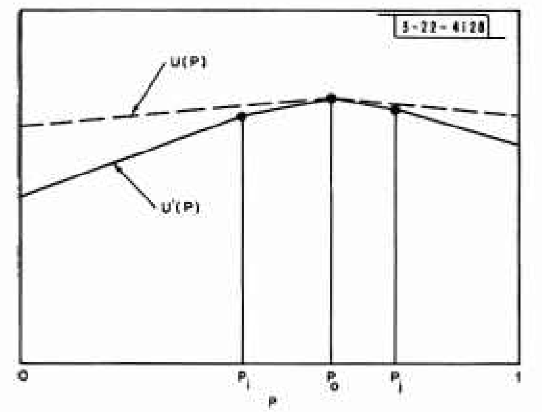


Fig. 7. Payoffs when μ is slightly greater than zero.

A pair of functions whose appearance satisfies the above properties is shown in Fig. 7. Here, the magnitude of the slope of $U(P)$ is equal to μ on either side of P_0. Now $U'(P)$ consists of four linear segments. As before, a breakpoint occurs at P_0, since the associated optimum look is different on either side of P_0. The point P_i is transformed into P_0 by an optimum look into box 2. It is a breakpoint of $U'(P)$ because the segments of $U'(P)$ to the left and right of P_i transform into the segments of $U(P)$ that are to the left and right of P_0, respectively. A similar effect occurs at P_j where $P_j \xrightarrow{1} P_0$.

If we continue to increase μ from zero on up, these functions will keep on changing in a continuous manner. The point P_0 may move, but P_i and P_j must be related to P_0 in the above manner. The two linear segments of $U(P)$ will continue to increase in steepness with μ, and so forth. Both functions must retain the same general appearance until at some μ_1 either the segment of $U(P)$ in

$(P_0, 1)$ becomes tangent to that of $U'(P)$ in (P_0, P_j), or the segment of $U(P)$ in $(0, P_0)$ becomes tangent to that of $U'(P)$ in (P_i, P_0).

Because the general appearance of the payoff functions and hence the general behavior of the associated optimum strategies for the two players are the same over the interval $(0, \mu_1)$, this interval is called a strategy interval. When μ increases beyond μ_1, the payoff functions take on new forms associated with the next strategy interval (μ_1, μ_2). We shall see that there is a sequence of strategy intervals $(0, \mu_1)$, (μ_1, μ_2), ..., (μ_k, μ_p) over $(0, \mu_p)$. The appearance of the payoff functions and the general behavior of the optimum strategies are the same over each interval but change from interval to interval. We shall see that as μ goes from interval to interval, these characteristics approach those associated with F^∞.

In order to extend this discussion in a more precise way, it is necessary to develop the properties of the payoff functions, which have already been discussed, more fully. Recall that both $U(P)$ and $U'(P)$ are continuous and convex, that $U(P)$ is identical to $U'(P)$ over the no-move region, defined as the interval in which

$$\left|\frac{dU'(P)}{dP}\right| < \mu \quad ,$$

and that the searcher's optimum strategy requires a look into box 2 if P is less than some unique P_0 and a look into box 1 if P is greater than P_0.

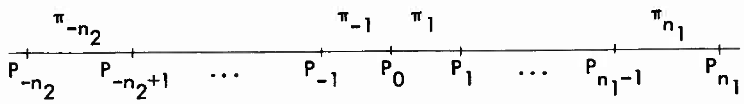
Let us consider the linear relationships that exist between $U(P)$ and $U'(P)$. The function $U'(P)$ is equal to $U'(P; 1)$ if $P \geqslant P_0$, and is equal to $U'(P; 2)$ if $P \leqslant P_0$. If $U(P)$ is linear over some interval π_k, then $U'(P; i)$ must be linear over the interval that is transformed into π_k by an unsuccessful look into box i. Therefore, if $U(P)$ is piecewise linear, $U'(P)$ must be also. $U(P)$ is identical to $U'(P)$ over the no-move region and is linear over each of the moving regions. Hence, the reverse is also true. We shall assume that both functions are indeed piecewise linear. That is, we shall assume that each function is partitioned into a set of linear segments by a set of breakpoints. As long as we can show that the set of breakpoints associated with each payoff function is finite, this assumption must be correct. We shall soon see that when $r_1^{n_1} = r_2^{n_2}$, $U(P)$ and $U'(P)$ are piecewise linear over the entire interval $(0, 1)$ as long as μ is finite. Also, we shall see that both functions are piecewise linear over the interval $(0, 1)$ in general if μ is strictly less than μ_p.

Let us consider the manner in which the breakpoints of the two functions are related to each other:

$$U'(P) = \min \begin{cases} U'(P; 1) = 1 + [Pr_1 + 1 - P]\, U\left[\dfrac{Pr_1}{Pr_1 + 1 - P}\right] \\[2ex] U'(P; 2) = 1 + [P + (1 - P)\, r_2]\, U\left[\dfrac{P}{P + (1 - P)\, r_2}\right] \end{cases} \quad .$$

The point P_0 must be a breakpoint of $U'(P)$ because it is the unique point at which $U'(P; 1)$ and $U'(P; 2)$ are equal. To either side of any other point, the same next look is optimum. Therefore, such a point can be a breakpoint of $U'(P)$ if and only if it is transformed by the next optimum look into a breakpoint of $U(P)$. $U(P)$ is linear over each of the moving regions. Therefore, all its breakpoints must belong to the no-move region where $U(P) = U'(P)$. It follows that any breakpoint of $U'(P)$ other than P_0 must be transformed into some other breakpoint of $U'(P)$ that belongs to the no-move region. Such a point must eventually be transformed into P_0 and it cannot be transformed into a moving region before this occurs. Therefore, the breakpoints other than P_0 that are common to U(P) and U'(P) are those points belonging to the no-move region that are transformed by an optimum search sequence into P_0 before leaving this region. The remaining breakpoints of U'(P) are those points transformed into a breakpoint of the no-move region by the next optimum look.

In F^{∞}, we found that the general behavior of the optimum search sequence when $r_1^{n_1} = r_2^{n_2}$ could be found by ordering the breakpoints in the recurrent region as follows:



We saw that any interval to the right of P_0 was transformed n_2 places to the left by an optimum look into box 1 and that any interval to the left of P_0 was transformed n_1 places to the right by an optimum look into box 2. A chain diagram could be drawn which would show the manner in which the linear intervals transformed into each other. If this was done, it was a fairly straightforward task to calculate the linear payoffs associated with each interval π_i and the values of the separating breakpoints.

In games F and F' a similar technique can be used. As we saw in the last section, the recurrent region (P_{01}, P_{02}) has the same properties that it has in F^{∞}. Although we may no longer equate P_0 to $q_2/(q_1 + q_2)$, we can still order the points $P_{-n_2}, \ldots, P_0, \ldots, P_{n_1}$ as before. These are the points belonging to the recurrent region that would be transformed into P_0 by an optimum search sequence if no moving were to occur. Therefore, these are the only points belonging to the recurrent region that can be breakpoints of U(P) or U'(P). If μ is less than μ_p, at least one of the moving regions must extend into the recurrent region. In this case, some of these points cannot be breakpoints of U(P) and usually some of them won't be breakpoints of U'(P) either.

As an example, consider again the case where $r_1^4 = r_2^3$ and suppose that the no-move region is (P_{-1}, P_3). It is a simple matter to find where the breakpoints occur. They are shown in Fig. 8.

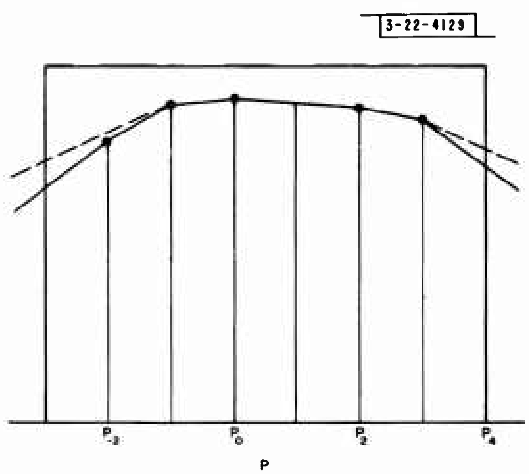


Fig. 8. The form of a possible pair of payoff functions.

Breakpoints of both U(P) and U'(P) occur at P_{-1}, P_0, P_2 and P_3. In addition, P_{-2}, which belongs to the recurrent but not the no-move region, is a breakpoint of U'(P). Breakpoints of U'(P) that lie outside the recurrent region are not shown.

This set of breakpoints is consistent. Each breakpoint of U'(P) to the left of P_0 transforms into a breakpoint in U(P) four (n_1) places to the right and so forth. Therefore, it is possible that the payoff functions may take on this form over some range in μ. A little thought will show, on the other hand, that if (P_{-2}, P_2) were guessed for the no-move region, inconsistencies would develop.

4.6 EXAMPLE: $r_1^4 = r_2^3$

It is appropriate at this point to return to our study of the manner in which the payoff functions behave as μ goes from zero to μ_p. In particular, let us again consider the example where $r_1^4 = r_2^3$. Figure 9 shows the various forms that these functions assume in the recurrent region (P_{-3}, P_4). The boundary points of the no-move region are denoted by circles. The linear intervals are numbered in order, $\pi_1, \pi_2, \ldots,$ and $\pi_{-1}, \pi_{-2}, \ldots,$ starting from P_0 and working toward P_- and P_+. Note that there is no longer any general correspondence between the subscript of an interval π_i and its bounding breakpoints P_j and P_k. The linear intervals of U(P) in the moving regions are designated by π_- and π_+, and $U_-(P)$ and $U_+(P)$ have slopes equal to μ and $-\mu$, respectively. The linear intervals of U'(P) that immediately adjoin (P_-, P_+) are designated by π'_- and π'_+.

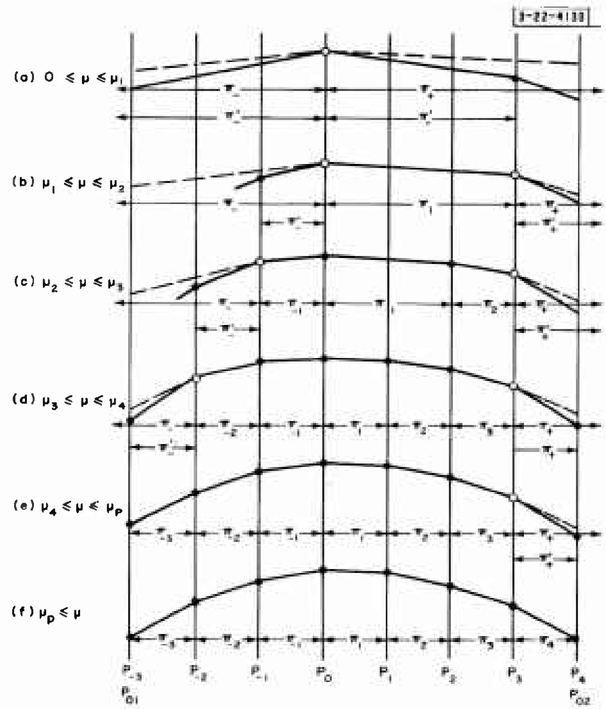


Fig. 9(a-f). The forms of the payoff functions for a set of strategy intervals $(r_1^4 = r_2^3)$.

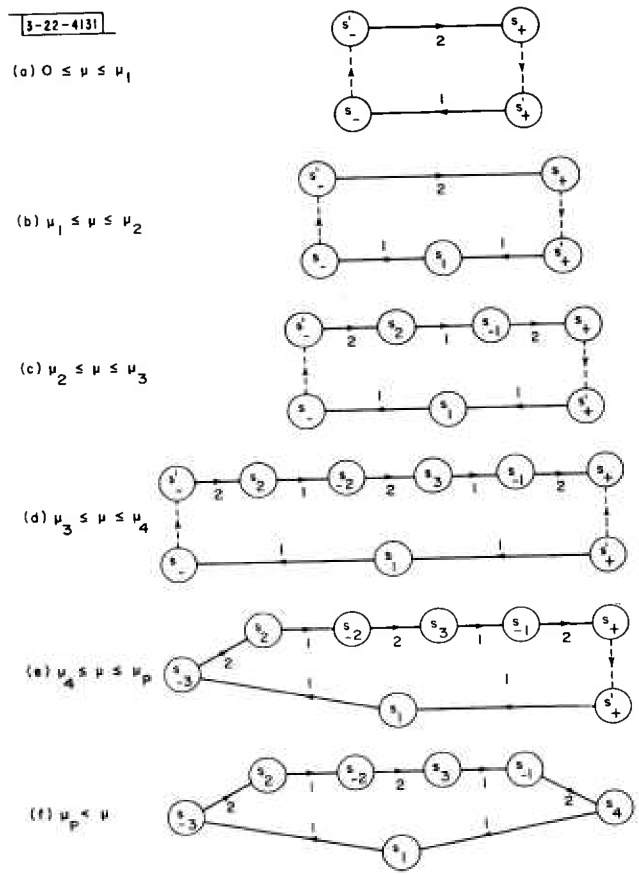


Fig. 10(a–f). Chain diagrams associated with the payoff functions of Fig. 9.

In Fig. 9(a), the general form of the payoff functions in the recurrent region is shown for μ close to zero. This result agrees with the discussion at the beginning of Sec. 4.5. As μ increases, the segments in π_- and π_+ increase in steepness. At some point, either P_- will shift to P_{-4} or P_+ will shift to P_3. In general, there is no simple way to predict which shift will occur. In the next section, where the actual computation of the payoff functions is considered, a method for computing which shift occurs first will be discussed. In this example, $U_+(P)$ becomes tangent to $U'_+(P)$ and P_+ shifts to P_3. The second strategy interval yields functions of the form shown in Fig. 9(b). Here a new breakpoint appears in $U'(P)$ at P_{-1} as a result of the breakpoint introduced at the new position of P_+.

As the moving cost increases further, we find the behavior exhibited in Fig. 9(c-e). In Fig. 9(e), P_- has shifted to the edge of the recurrent region. When $U_+(P)$ becomes tangent to $U'_+(P)$, μ is equal to μ_p. Figure 9(f) shows the general form when the moving cost is prohibitive. Here, of course, $P_0 = q_2/(q_1 + q_2)$.

If both q_1 and q_2 are unequal to one, the magnitude of the slope of $U^\infty(P)$ becomes arbitrarily large as P approaches zero and one. Therefore, as long as μ is finite, P_- must be greater than zero and P_+ must be less than one. It follows that U(P) and U'(P) can have only a finite number of breakpoints in the no-move region when $r_1^{n_1} = r_2^{n_2}$. Furthermore U'(P) can have only a finite number of breakpoints in either moving region. Therefore, when $r_1^{n_1} = r_2^{n_2}$, then U(P) and U'(P) will be piecewise linear over the interval (0, 1) if μ is finite.

Chain diagrams that illustrate the behavior of the optimum strategies can be drawn in a manner quite similar to that used in Chapter 2. Those associated with the various strategy intervals of our example are shown in Fig. 10. In these chain diagrams, each state s_i is associated with the interval π_i and has an associated payoff function. A transition from one state to another produced by an optimum look is represented by a solid line. The transitions from state s_- to s'_- and from s_+ to s'_+ occur when the evader moves (with the proper probability) and are represented by broken lines. In general, each linear interval that belongs to both the no-move and the recurrent regions will be represented by a single state s_i in the chain diagram since $U_i(P)$ and $U'_i(P)$ are identical. In addition to these states, s_- and s'_- will be included in the chain if P_- belongs to the interior of the recurrent region and s_+ and s'_+ will be included if P_+ does. Here we must differentiate between the states associated with U(P) and U'(P) since these payoffs are different.

In the case where $q_2 = 1$ and $q_1 \neq 1$, the behavior of the payoff functions is quite similar to that found in the above example. In this case, the breakpoints occur at $P_0, P_1, \ldots, P_k, \ldots, P_+$, where P_k is transformed into P_0 by k looks into box 1. There is one linear segment over $(0, P_0)$. In the first strategy interval, $P_- = P_0 = P_+$. As μ increases, P_+ will shift from P_0 to P_1 to P_2 and so forth. At some point, P_- must shift from P_0 to zero, and as μ increases further, P_+ will continue to shift to the right, point by point. Since P_k approaches $P_{02} = 1$ only as k approaches infinity, this process will continue indefinitely. Over any strategy interval, the chain diagram can be drawn in the same manner as in the example where $r_1^4 = r_2^3$.

In Chapter 2, we found that a chain diagram could not be associated with the linear segments of $U^\infty(P)$ in the recurrent region when $\log r_2/\log r_1$ was irrational because there were an infinite number of segments. That is, there were an infinite number of points in (P_{01}, P_{02}) that were eventually transformed into P_0 by the optimum search sequence. In games F and F' this no longer occurs as long as μ is strictly less than μ_p. In this event, at least one of the moving regions must extend into the recurrent region, and only a finite number of points can be transformed into P_0 by an optimum sequence without leaving the no-move region. As a result, these points partition the payoff functions U(P) and U'(P) into a finite number of linear segments and both functions are piecewise linear. A finite chain diagram that illustrates the manner in which these segments transform into each other can be drawn.

A consequence of this result is that an n_1/n_2 approximation of $\log r_2/\log r_1$ may be used to obtain an exact solution of the payoff function when $\mu < \mu_p$. The reason that an exact solution can be obtained by assuming a good choice of n_1 and n_2 follows from the fact that the actual computations of the payoff functions depend only on the chain diagram used. If the correct chain diagram is found, the resulting solution will be correct. The n_1/n_2 approximation can be used as a device for generating a sequence of chain diagrams. As long as the approximation is sufficiently accurate, it will produce a sequence of chains as μ increases that will agree with the sequence associated with the irrational case up to the correct one. The former sequence will merely be finite whereas the latter is infinite. Clearly, the approximation must be increasingly more accurate and the resulting chain diagram will become increasingly large as μ approaches μ_p.

As an example, let us approximate $n_1/n_2 = 4/3$ by $n_1'/n_2' = 3/2$. In this case, the 3/2 approximation will yield a sequence of three chain diagrams identical to the first three in Fig. 10. If μ is sufficiently small, one of these three will be the correct one and the correct solution can be obtained.

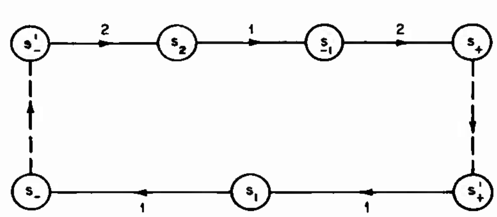
4.7 COMPUTATION OF THE PAYOFF FUNCTIONS

The computation of the payoff functions is accomplished in two steps. First, the correct chain diagram must be found. Once this has been done, the payoff functions $U_i(P)$ and $U_i'(P)$ associated with each interval π_i that has a corresponding state s_i in the chain can be calculated. Finally, the separating breakpoints can be found. In the last section we saw that the chain diagrams changed from one strategy interval to another and that at the end of each strategy interval two possible changes could occur. In this section we shall see how the correct change can be determined. The required computations, although simpler, are quite similar to those used to compute the actual payoff functions within a strategy interval. Therefore, we shall consider the latter problem first.

4.7.1 Computation of the Payoff Functions When the Correct Chain Diagram Is Known

The chain diagram associated with a given strategy interval contains all the information needed for computing the payoff functions. In order to clarify the discussion, the chain diagram in Fig. 10(c) will be used as an example.

3-22-4150



This diagram is typical of those that occur when both P_- and P_+ belong to the interior of the recurrent region. Each linear interval belonging to the no-move region, where $U(P) = U'(P)$, has a corresponding state in the chain. The states associated with π_-, π'_-, π_+ and π'_+ are also included, since these intervals extend into the recurrent region. In such a chain there is a single loop. Moving occurs only in the transition from s_- to s'_- and from s_+ to s'_+. These two pairs of states divide the loop into two parts.

The linear payoffs associated with each state can be expressed in the same form used in F^∞. For any state s_i associated with an interval in the no-move region we can write

$$U_i(P) = U_i'(P) = a_iP + b_i(1 - P) \quad .$$

Furthermore, we can let

$$U_-(P) = a_-P + b_-(1 - P) \quad ,$$

$$U'_-(P) = a'_-P + b'_-(1 - P) \quad ,$$

and so forth.

If two states are not separated from each other by a move transition, their payoffs are related to each other in the same manner as in F^∞. If a look sequence represented by $\{t_m(n)\}$ transforms s_i into s_j and no move transitions intervene, we may use Eq. (2-13) to write

$$a_i = q_1 \sum_{n=1}^{k_1} t_1(n)\, r_1^{n-1} + r_1^{k_1}(k_1 + k_2 + a_j) \quad ,$$

$$b_i = q_2 \sum_{n=1}^{k_2} t_2(n)\, r_2^{n-1} + r_2^{k_2}(k_1 + k_2 + b_j) \quad . \qquad (4\text{-}8)$$

As before, k_i represents the total number of looks into box i during the transition. Since at least one move transition occurs in the chain when μ is less than μ_p, the payoff associated with a given state can never be expressed in terms of itself and solved directly. The above equations can be used, however, to express $U'_-(P)$ in terms of $U_+(P)$ and $U'_+(P)$ in terms of $U_-(P)$. This will yield four equations involving the eight unknowns a_-, b_-, a'_-, b'_-; a_+, b_+, a'_+ and b'_+.

Other properties that can be utilized in order to get a complete set of equations are as follows. First, the magnitudes of the slopes of $U_-(P)$ and $U_+(P)$ are equal to μ. Therefore,

$$a_- - b_- = \mu \quad ,$$

$$b_+ - a_+ = \mu \quad . \qquad (4\text{-}9)$$

Also, $U_-(P)$ and $U'_-(P)$ must intersect at P_-, and $U_+(P)$ and $U'_+(P)$ must intersect at P_+. This yields the equations

$$a_-P_- + b_-(1 - P_-) = a'_-P_- + b'_-(1 - P_-) \quad ,$$

$$a_+P_+ + b_+(1 - P_+) = a'_+P_+ + b'_+(1 - P_+) \quad , \qquad (4\text{-}10)$$

bringing the total to eight equations. With the addition of the unknown P_- and P_+, however, there are now ten unknowns.

These points are the bounding points of the no-move region. They are related to P_0 by the manner in which they are transformed into it by the optimum search sequence. These sequences can be found by looking at the chain diagram. In particular, $P_- \to P_0$ as $s'_- \to s_{-1}$, and $P_+ \to P_0$ as $s'_+ \to s_{+1}$. No moving transitions occur during the above processes since the states in the chain diagram form a single loop in which

$$s_{-1} \xrightarrow{2} s_+ \quad \text{and} \quad s_{+1} \xrightarrow{1} s_- \quad .$$

As before,

$$P_i \xrightarrow{(k_1,\, k_2)} P_j \Longrightarrow P_i = \frac{P_j r_2^{k_2}}{P_j r_2^{k_2} + (1 - P_j)\, r_1^{k_1}} \quad . \tag{4-11}$$

The two equations of this form introduce the additional unknown P_0 and one more equation is necessary to complete the set. The final property which can be used is that a look into either box is optimum when P is equal to P_0. Since P_{01} belongs to π_- and P_{02} belongs to π_+,

$$\begin{aligned} U(P_0) &= 1 + [P_0 r_1 + 1 - P_0]\, U_- \left[\frac{P_0 r_1}{P_0 r_1 + 1 - P_0}\right] \\ &= 1 + [P_0 + (1 - P_0)\, r_2]\, U_+ \left[\frac{P_0}{P_0 + (1 - P_0)\, r_2}\right] \quad , \end{aligned}$$

which reduces to the equation

$$a_- P_0 r_1 + b_-(1 - P_0) = a_+ P_0 + b_+(1 - P_0)\, r_2 \quad . \tag{4-12}$$

This completes the set of equations from which a solution can be obtained. It should be noted that the number of states in the chain has no effect on the number of equations required, and these equations apply whenever both P_- and P_+ belong to the interior of the recurrent region. Most of the above equations are linear and express only one unknown in terms of another. The complete set can be reduced to a single cubic equation in a fairly direct manner. It is usually most convenient to derive this cubic equation in terms of P_0. Once this has been done, the other variables in the set of equations can be obtained easily. Finally, the payoff functions associated with the other states in the chain and the remaining breakpoints can be calculated by using the same techniques used in Chapter 2.

To illustrate the general method, let us write the equations appropriate to the chain diagram at the beginning of this section. We see that $s'_- \xrightarrow{212} s_+$ and $s'_+ \xrightarrow{11} s_-$. From Eq. (4-8) it follows that

$$a'_- = q_1(2) + r_1(3 + a_+) \quad ,$$

$$b'_- = q_2(1 + 3r_2) + r_2^2(3 + b_+) \quad ,$$

$$a'_+ = q_1(1 + r_1) + r_1^2(2 + a_-) \quad ,$$

$$b'_+ = 2 + b_- \quad .$$

Equations (4-9), (4-10) and (4-12) may be used directly. Finally, since $P_+ \xrightarrow{1} P_0$ and $P_- \xrightarrow{21} P_0$, we may write

$$P_+ = \frac{P_0}{P_0 + (1 - P_0)\, r_1}$$

and

$$P_- = \frac{P_0 r_2}{P_0 r_2 + (1 - P_0)\, r_1} \quad .$$

These equations and the chain diagram are appropriate for the case where $r_1 = 0.512$ and $r_2 = 0.4096$ ($r_1^4 = r_2^3$) when μ is equal to 1.3. These are the same escape probabilities used in the example in Chapter 2. The payoff functions which result are presented in Table I and graphed in Fig. 11. The quantity U(P) is a maximum at $P_0 = 0.528$ and is equal to 3.243. Therefore, the evader should initially hide in box 1 with this probability and can guarantee a payoff equal to 3.243. The quantities P_- and P_+ are equal to 0.482 and 0.694, respectively, and define the no-move region. With these, the evader's moving strategy is easily calculated. If P is less than P_-, he should move to box 1 if in box 2 with probability $x_1 = (0.482 - P)/(1 - P)$. If P is greater than P_+, he should move to box 2 if in box 1 with probability $x_2 = (P - 0.694)/P$.

It is worth noting that the equations used to compute these functions did not make use of the fact that $r_1^4 = r_2^3$. The solution is correct because the correct chain diagram was used and no contradictions occurred. The contradictions that would arise if the wrong diagram were used are quite simple: either the magnitude of the slope of U'(P) would be less than μ in a moving region (in π'_- of π'_+) or it would exceed μ somewhere inside the no-move region. The slope of each linear segment in π_i is equal to $a_i - b_i$ and is included in Table I.

If only P_+ or P_- belongs to the interior of the recurrent region, the solution is somewhat simpler. As an example, consider the chain diagram in Fig. 10(e). Here only P_+ belongs to the interior of the recurrent region, and neither s_- nor s'_- occurs in the chain. As a result, a'_+ and b'_+ may be expressed in terms of a_+ and b_+ by means of Eq. (4-8). None of the equations that involve P_-, a_-, b_-, a'_- and b'_- are required. On the other hand, Eq. (4-12) must be rewritten and some new variables must be introduced into the set of equations. Previously, a_- and b_- were included in this equation because P_{01} belonged to the interval π_-. In this example, P_{01} is the breakpoint separating π_- and π_{-3}. Since s_{-3} is included in the chain diagram, while s_- is not, it is worthwhile to rewrite Eq. (4-12) in the form

$$a_{-3} P_0 r_1 + b_{-3}(1 - P_0) = a_+ P_0 + b_+(1 - P_0)\, r_2 \quad .$$

We can express a_{-3} and b_{-3} in terms of a_+ and b_+, respectively, by means of Eq. (4-8). The set of equations that results can be reduced to a single quadratic equation in P_0. In other respects, the solution is accomplished in the same manner as before.

4.7.2 Determination of the Correct Chain Diagram

Unless one wishes to guess the form of the chain diagram that applies for a given pair of boxes and a particular moving cost, one can examine the manner in which the form of the payoff functions changes from one strategy interval to another. Two problems should be apparent. First, one must find which of the two possible changes occurs when μ moves from one strategy

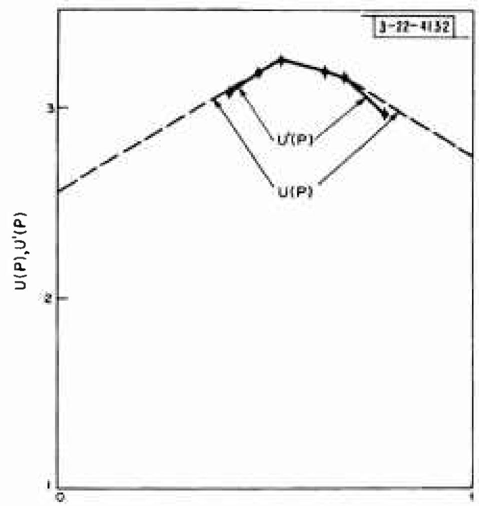


Fig. 11. Payoff functions ($r_1 = 0.512$, $r_2 = 0.4096$, $\mu = 1.3$).

TABLE I

PAYOFF FUNCTIONS

($r_1 = 0.512$, $r_2 = 0.4096$, $\mu = 1.3$)

Payoff Function	Range	Slope ($a_i - b_i$)
$U'_-(P) = 3.918P + 2.498(1-P)$	(0.427, 0.482)	1.420
$U_-(P) = 3.856P + 2.556(1-P)$	(0, 0.482)	1.3
$U_{-1}(P) = 3.747P + 2.658(1-P)$	(0.482, 0.538)	1.089
$U_1(P) = 2.974P + 3.556(1-P)$	(0.538, 0.645)	−0.582
$U_2(P) = 2.918P + 3.658(1-P)$	(0.645, 0.694)	−0.739
$U_+(P) = 2.747P + 4.047(1-P)$	(0.694, 1.0)	−1.3
$U'_+(P) = 2.523P + 4.556(1-P)$	(0.694, 0.780)	−2.033

Note: $P^* = P_0 = 0.538$

$U(P^*) = 3.243$

interval to another. In addition, one must determine the point at which the change occurs, i.e., the value of the moving cost at the change point.

In order to illustrate a method that can be used to answer these questions, let us consider the manner in which the third strategy interval (μ_2, μ_3) of Fig. 9 changes into the fourth. We know that the third strategy interval is correct over some range of μ when $r_1 = 0.512$ and $r_2 = 0.4096$ since it gave a valid solution for $\mu = 1.3$. To find which change occurs at μ_3, we must make a guess and find if it is correct. For convenience, let us make the right one. That is, let us assume that $U'_-(P)$ becomes tangent to $U_-(P)$ as μ approaches μ_3.

When μ is exactly equal to μ_3, then $U'_-(P)$ must be identical to $U_-(P)$. Therefore, in the chain diagram in Fig. 10(c) we can delete the dotted line joining s_- and s'_-, and we can express $U'_+(P)$ in terms of $U_+(P)$. However, μ_3 must be left as an unknown.

The equations that result in this particular example are as follows. Since $s'_+ \xrightarrow{212} s_- = s'_- \xrightarrow{11} s_+$, we may write

$$a'_+ = q_1(2 + 4r_1 + 5r_1^2) + r_1^3(5 + a_+) \quad ,$$

$$b'_+ = q_2(1 + 3r_2) + r_2^2(5 + b_+) \quad ,$$

$$a'_- = a_- = q_1(1 + r_1) + r_1^2(2 + a_+) \quad ,$$

$$b'_- = b_- = 2 + b_+ \quad .$$

No equations involving P_- are required and, in fact, P_- is not unique. On the other hand, $U_+(P)$ is equal to $U'_+(P)$ only at P_+. Therefore,

$$a_+P_+ + b_+(1 - P_+) = a'_+P_+ + b'_+(1 - P_+) \quad ,$$

where $P_+ \xrightarrow{1} P_0$, or

$$P_+ = \frac{P_0}{P_0 + (1 - P_0)\, r_1} \quad .$$

Again Eq. (4-12) must be used, and Eq. (4-9) may be expressed in the form

$$b_- - a_- = a_+ - b_+ = \mu_3 \quad .$$

The solution reveals that $\mu_3 = 1.393$. The magnitude of the slope of s'_+ proves to be equal to 1.973. Since this value is greater than μ_3, no contradiction arises and our guess was correct. The form of the payoff functions in the fourth strategy interval must therefore be that shown in Fig. 9(d).

The values of $\mu_0, \mu_1, \mu_2, \ldots, \mu_p$, given in Table II, indicate the range over which each of the chain diagrams of Fig. 10 is valid when $r_1 = 0.512$ and $r_2 = 0.4096$. The maximum payoff $U(P^*)$ is also included for each μ_i. These payoffs indicate the manner in which the value of the game decreases as μ increases from zero to μ_p. Note that as μ gets close to μ_p the value decreases very slowly.

TABLE II

THE BOUNDARIES OF THE STRATEGY INTERVALS AND THE VALUE AT EACH BOUNDARY ($r_1 = 0.512$, $r_2 = 0.4096$)

i	μ_i	$U(P^*)$
0	0	3.743
1	0.903	3.297
2	1.170	3.256
3	1.393	3.234
4	1.541	3.227
p	1.913	3.222

CHAPTER 5
THE SEARCHER'S GOOD STRATEGY

5.1 INTRODUCTION

In the last chapter, the evader's good strategy was developed by assuming that his optimum strategy in the modified game was indeed his good strategy in G. In the process, we found that he could guarantee a payoff equal to U(P) if he selected P initially, and that he could guarantee a maximum payoff U(P*). The searcher's optimum strategy, which limited the evader to the above payoff in the modified game, proved to be quite similar to that in F^{∞}. In fact, we were able to solve G completely when μ was prohibitive because the game degenerated to a form effectively the same as that of G^{∞}.

In this chapter, we shall develop the searcher's good strategy in G when the moving cost is not prohibitive. It will be shown that the searcher can limit the evader to U(P) if he knows only the initial P that the evader selects and no more. This statement applies even if the evader knows the strategy used by the searcher. Once the evader has been thus limited, the solution can be extended to G, where even the initial P is unknown, in much the same way as it was in G^{∞}. This good strategy will limit the evader to the payoff U(P*).

The actual computation of the searcher's good strategy will prove to be fairly easy because this good strategy is strongly related to the function U(P) and the chain diagram utilized in computing it. Most of the work has been done once U(P) has been found. As we shall see, the searcher's good strategy will be Markovian in form. Therefore, it is appropriate to examine some basic properties of Markovian search strategies before considering the relationship between the searcher's good strategy in G and his optimum strategy in F.

5.2 MARKOVIAN SEARCH STRATEGIES AND MODIFIED GAMES H AND H'

In this section we shall consider search strategies that generate a search sequence by means of a discrete-time Markov process. Such a process is a mathematical model defined by a set of states, a set of transitions between these states, and an associated set of transition probabilities. Given a particular state, a transition will occur in the next time interval to some other state, or possibly to the same state, according to the set of transition probabilities associated with that state. A search sequence can be generated by such a process if a particular look is associated with each transition and a probability distribution for selecting a starting state is defined.

Discrete-time Markov processes have most often been used to model the behavior of a physical system. In such a situation, each state is defined by a particular set of values for a set of variables that completely characterize the system at any given time. As a result, the primary interest usually focuses on these states. For example, one may wish to calculate the probability that the system will be in a particular state after k units of time if it is originally in a known state.

When Markov search strategies are considered, however, the primary interest shifts to the looks and hence to the transitions, for the Markov process is used strictly as a device for generating a search sequence. Any discrete-time Markov process may be used once a look is associated with each transition, and we need not consider the physical significance of any state. As we shall find in Sec. 5.5, each state in the process defined by the searcher's good strategy will have some significance. It is not appropriate at this point, however, to concern ourselves with this problem. The only restrictions which will be imposed on the Markov processes are that the number of states must be finite and that, at most, one transition from each state can be associated with a given look. Since there are only two boxes, there can, of course, be at most two transitions from each state.

A simple example of a Markovian search strategy that obeys these constraints is defined by a transition diagram (Fig. 12), a set of transition probabilities and the probability distribution $Y_0 = \{y_0(\sigma_1), y_0(\sigma_2), y_0(\sigma_3), y_0(\sigma_4)\}$, which is used to select the starting state. In contrast to

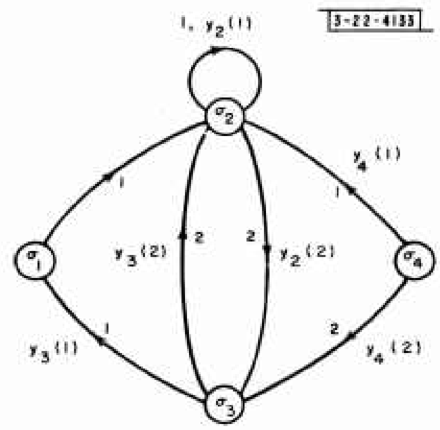


Fig. 12. The transition diagram of a Markovian search strategy.

the usual convention in which p_{ij} is used to represent the probability of a transition to σ_j, given σ_i, here $y_i(k)$ is used to represent the probability that box k is examined next, given σ_i. The term $\sigma_{i|k}$ will be used to represent the state that follows when this event occurs.

The above transition diagram exhibits several properties worth noting. First, σ_4 can be occupied only at the beginning of the process, since after the first look no transitions can be made into it. Therefore σ_4 is a special example of a transient state. In general, a state will be a transient state if the probability that it can be occupied approaches zero as the process continues indefinitely. Clearly, σ_1, σ_2 and σ_3 are not transient states and, in fact, belong to a single recurrent chain. A recurrent chain consists of a set of states in which it is always possible to get from one to any other by a series of transitions. Once a state belonging to a recurrent chain is entered, only states belonging to that chain can be occupied in the future. Furthermore, once this has occurred, the probability that each of the states in the chain is occupied as the process continues indefinitely approaches a nonzero limiting value. It will develop that Markov processes with only one recurrent chain will be sufficient in our study.

The final property which we should note is that only one transition can occur from σ_1 and that the next look associated with this state is deterministic. A state of this type will be called a pure state. States from which more than one transition is possible will be called mixed states.

In order to discuss the influence that such a Markovian search strategy has on the behavior of the search evasion game, it is helpful to introduce the modified games H and H'. These games are similar to the modified games F and F', but here we reverse things and require the searcher to reveal part of his search strategy to the evader. In particular, he must reveal the transition diagram and the associated transition probabilities that he uses and must tell the evader which state is initially selected once the evader has hidden. He is not required, however, to reveal the probability distribution Y_0 used for this selection, and for the time being we shall not concern ourselves with it. In the same manner as before, H applies when the evader still has an opportunity to move before the next look, and H' applies after this opportunity has passed.

Payoff functions can be associated with each of these games, but now a different pair must always be associated with each state σ_i since both players are aware of the state that applies at any given time. The quantity $W_i(P)$ will be used to represent the future payoff that applies in H if the search process is in σ_i and if the evader is in box 1 with probability P and uses an optimum strategy in the future. The quantity $W_i'(P)$ will be used to represent the corresponding payoff in H'. No statement concerning the searcher's future strategy is included in these definitions, since it is completely specified by the Markov process.

Although the searcher is no longer informed of the value of P that applies at any given time and the evader always knows exactly where he is, these payoffs are still functions of P. This variable is the one that an observer would use to define the evader's position if he knew both players' strategies and was able to observe the search sequence which resulted. This assumes, of course, that the observer cannot see when the evader actually moves. In these games, the evader's moving strategy may now be a function of the search state as well as a function of P and his own position.

A pair of functional equations can be written to express the payoffs associated with H in terms of those associated with H' and vice versa. In H', the searcher's next look is completely specified by the Markov process. Given σ_i, he will look into box 1 with probability $y_i(1)$ and into box 2 with probability $y_i(2)$. Therefore,

$$W_i'(P) = 1 + y_i(1)[Pr_1 + 1 - P]\, W_{i|1}\left[\frac{Pr_1}{Pr_1 + 1 - P}\right]$$

$$+ y_i(2)[P + (1-P)\, r_2]\, W_{i|2}\left[\frac{P}{P + (1-P)\, r_2}\right] \quad . \qquad (5\text{-}1)$$

Here, $W_{i|k}(P)$ represents the payoff in H associated with the state that follows σ_i if box k is examined. If σ_i is a pure state, the above equation will of course degenerate to a simpler form.

In H, the evader has the opportunity to move. As before, the cost function $C(P \rightarrow P')$ is associated with a transformation of the state variable. Since the evader can calculate the payoffs $\{W_i'(P)\}$,

$$W_i(P) = \max_{P'} \{-\mu|P - P'| + W_i'(P')\} \quad .$$

Each payoff function $W_i'(P)$ must be linear and is valid for all P in (0, 1). It follows that

$$W_i(P) = \begin{cases} -\mu P + W_i'(0) \quad , & \dfrac{dW_i'(P)}{dP} < -\mu \\[2ex] W_i'(P) \quad , & -\mu \leqslant \dfrac{dW_i'(P)}{dP} \leqslant \mu \\[2ex] -\mu(1-P) + W_i'(1) \quad , & \dfrac{dW_i'(P)}{dP} > \mu \quad . \end{cases} \qquad (5\text{-}2)$$

When the slope of $W_i'(P)$ is greater than μ, the evader must move to box 1 if he is in box 2, and when the slope is less than $-\mu$ he must move to box 2 if he is in box 1. On the other hand, if $|dW_i'(P)/dP|$ is strictly less than μ, the evader should not move, and $W_i(P)$ is equal to $W_i'(P)$. $W_i(P)$ is also identical to $W_i'(P)$ when $|dW_i'(P)/dP| = \mu$, but here the evader can still move. If $dW_i'(P)/dP = \mu$, the evader can move to box 1 if he is in box 2 with any probability and can therefore increase P by any desired amount. The reverse holds when $dW_i'(P)/dP = -\mu$. This property is very important, for it allows the evader's optimum strategy in F to be consistent with his optimum strategy here when the searcher uses the correct Markov process. The three possible ways in which the payoffs $W_i(P)$ and $W'_i(P)$ can be related to each other are illustrated in Fig. 13. Here, $W_i(P)$ is indicated by a broken line if it is unequal to $W_i'(P)$.

Once the Markov process, except for the starting rule Y_0, is specified, the evader's optimum strategy may be obtained by using a form of linear programming. Such a solution will maximize $W_i(P)$ for all P in each σ_i. Although we do not need to concern ourselves with the manner in which such a solution can be obtained, it is worthwhile to discuss some properties implied by the result.

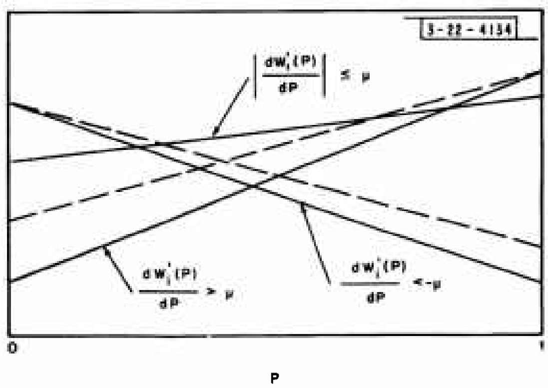


Fig. 13. The relation between $W_i(P)$ and $W_i'(P)$.

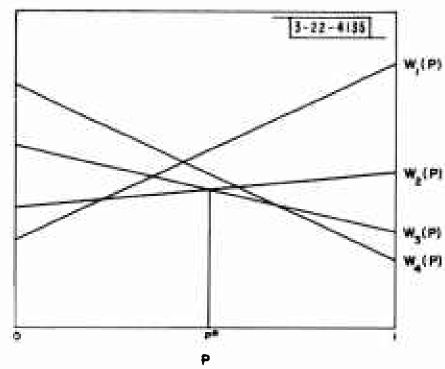


Fig. 14. A possible set of payoff functions for the strategy shown in Fig. 12.

Let us suppose that the Markov process of Fig. 12 yields the solution shown in Fig. 14. Let us assume that $dW_4'(P)/dP = -\mu$ and that $dW_1'(P)/dP > \mu$. In both σ_1 and σ_4, $dW_i(P)/dP = \mu$. In σ_1 however, the evader must move to box 1 if he is in box 2, whereas in σ_4 he can decrease P by any desired amount. In σ_2 and σ_3, on the other hand, he should not move. Note that in general $|dW_i(P)/dP| \leqslant \mu$. As long as Y_0 is unknown to the evader, he should initially hide in box 1 with probability P* since this guarantees the maximum payoff of $W_2(P^*) = W_3(P^*)$. Of course, if the searcher started the Markov process in σ_1 or σ_4 the evader would receive more. The searcher would be foolish to do this, however, for there exists a $Y_0 = (0, y_0(\sigma_2), y_0(\sigma_3), 0)$ which limits the evader to the above amount.

Unfortunately, such a solution does not guarantee that this is the searcher's good strategy in G, for we have no reason to assume that the transition probabilities or the transition diagram is correct. We can be sure that such a strategy is the good strategy only if it limits the evader to U(P*). Clearly, it would be a formidable task to guess the transition probabilities, let alone the transition diagram associated with the good search strategy, if we did not have some guidelines to help us on our way. For this reason the evader's good strategy has been developed first.

Before we consider how the good Markovian search strategy is related to the behavior of game F, however, a fundamental property of good strategies should first be discussed.

5.3 AN IMPORTANT PROPERTY OF GOOD STRATEGIES

In general, a two-person zero-sum game has a value, and a pair of good strategies exists, if each player can guarantee that he will receive a payoff no worse than the value. The strategy that guarantees this payoff is the player's good strategy, and each good strategy is optimum against the other. If one player tells the other that he is using his good strategy, the other player can gain no advantage by using a strategy different from his good strategy.

This behavior is quite different from that associated with any other pair of strategies. If one player were to use an arbitrary one and inform the other of what it was, the other player could also use a different strategy and collect a larger payoff. If in turn, he told the first player what this new strategy was, that player would probably decide to use a different one himself. This process can be continued and leads to a "if I do this, he will do thus and so, but then I should do something else, but then he will. . . . " type of reasoning. Only the good strategies avoid instabilities of this type.

In most games of interest (excluding perfect information games such as chess) each player's good strategy involves random decisions. Such a strategy is called a mixed strategy if the game is expressed in normal form. On the other hand, it can be expressed in terms of a set of behavioral strategies as we shall do here. As was mentioned earlier, a behavioral strategy associates with each information set or behavioral state for the player in question a probability distribution for selecting the next alternative. In general, the probability distribution associated with a given behavioral state need not include a nonzero probability for each alternative. An alternative that does have a nonzero probability in a given behavioral state can be called an admissible alternative of that state. Alternatives that occur with probability zero will be called inadmissible alternatives of that state. This of course holds only when the number of alternatives in each state is finite as it is in the search evasion game.

The property of the good strategies that we wish to discuss here is as follows. If one player uses his good strategy, the payoff will be equal to the value of the game as long as the other player selects only admissible alternatives. That is, the payoff is the same for any set of probability distributions over the behavioral states of one player as long as these distributions exclude the selection of inadmissible alternatives <u>and</u> the other player uses his good strategy.

As an example, consider the good strategies in G^{∞}. The evader's good strategy requires him to hide in box 1 with probability P^* and in box 2 with probability $1 - P^*$. Thus, hiding in either box is admissible. The searcher's good strategy, on the other hand, requires him to choose one of the two infinite search sequences optimum at P^*, and these two sequences are his admissible alternatives. As long as the evader uses his good strategy P^*, the payoff will equal the value if either of these two sequences is selected. Similarly, the probability distribution that the searcher uses to choose one of these sequences causes the payoff to be independent of P and, therefore, equal to the value for either of the evader's admissible alternatives. Note that in this game the evader has no inadmissible alternatives, whereas the searcher has an infinite number.

This property of good strategies is very useful when one wishes to derive the good strategy for one player once the other's is known. Any alternative that causes the payoff to be unequal to the value when the other player uses his good strategy must be an inadmissible alternative for that behavioral state and can be excluded from consideration. Once this has been done, the problem of finding the "good" probability distributions over the admissible alternatives in each state is much simpler.

In the next section, we shall find that this property allows us to derive the transition diagram associated with the searcher's good strategy in a straightforward manner. The associated transition probabilities can then be computed by making use of the previously calculated function U(P). Finally, the initial distribution Y_0 can be found to complete the solution.

5.4 DERIVATION OF THE SEARCHER'S TRANSITION DIAGRAM

When a Markov process is used to generate a search sequence, each state in the transition diagram is a behavioral state of the searcher's strategy and each transition represents an alternative. To start this process, one of these states must be selected by means of Y_0. A starting state σ_0, not shown in the transition diagram, may be associated with this distribution. As would be expected, only some of the states in the transition diagram should be initially selected with a nonzero probability. The selection of these states corresponds to the admissible alternatives in σ_0. The general form of the transition diagram and also the admissible alternatives associated with σ_0 can be found by considering the behavior of game G when the evader is required to use his good strategy. It will be more convenient, however, to derive the form of the transition diagram first, and consider the start-up state σ_0 later.

In order to do this, we must modify slightly our restrictions on the evader's strategy. In G, the evader's good strategy contains two parts. First, he must use P* to determine where he hides initially, and then he must exercise his good moving strategy as the game is played. We can simplify things by assuming that the initial P is arbitrarily assigned and known to the searcher. Once it has been used, the evader is required to exercise his good strategy and, in fact, must move before the first look, if necessary. Under these conditions, the searcher can utilize the initial P in starting the search process, and we must find a transition diagram with which he can limit the evader to U(P).

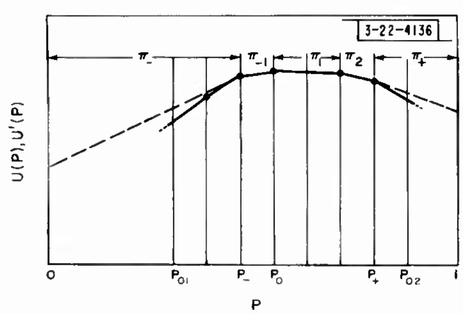
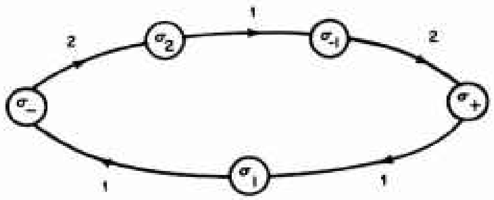


Fig. 15. A pair of payoff functions for F and F'.

In order to clarify the discussion, let us consider Fig. 15 [the payoff functions in Fig. 9(c)]. Let us also reproduce the associated chain diagram in Fig. 10(c) with the moving transitions eliminated.

If the evader were to use his good moving strategy, this diagram could be used to generate a search sequence that would yield a payoff equal to U(P). The process should merely be started in the state σ_i that corresponds to the interval π_i in which the assigned P lies. It follows that the look associated with each state in the chain must be admissible. Unfortunately, each state

3-22-4151



is a pure state and the resulting search sequence would be deterministic. Under these conditions, the evader could obviously secure a larger payoff by using a moving strategy other than his good one. Clearly, we have not found all of the searcher's admissible alternatives.

In order to determine when other looks are admissible, the evader's good moving strategy must be examined more closely. After a transition to σ_+ has occurred, P belongs to π_+ and the evader must transform it to P_+, the left boundary of π_+. The state σ_+ requires a look into box 1 and a transition to σ_1 occurs. In the process, P_+ is transformed to P_0, the left boundary of π_1. As a result, a look into either box is admissible, and if box 2 is examined, P will return to π_+. It follows that a look into either box is admissible in σ_1 once the search process has occupied σ_+. A look into box 2 produces a transition to σ_+, whereas a look into box 1 yields a transition to σ_- as before. Similar reasoning shows that a look into either box is admissible and the same transitions occur if the Markov process is in σ_{-1} and has occupied σ_- beforehand. Since this reasoning does not apply until σ_+ or σ_- has been occupied, we must differentiate between these two situations.

The transition diagram that takes this into account is shown in Fig. 16. In this diagram, the states σ_i^t and σ_i^r are associated with each interval π_i in the no-move region. State σ_i^t is transient and applies before a move occurs. State σ_i^r belongs to the recurrent chain and applies thereafter. States σ_- and σ_+ also belong to the recurrent chain but have no superscripts because there are no corresponding transient states. These will be called the moving states, since the evader's good moving strategy requires a transformation of the state variable P in each one of them. If P belongs initially to π_i, the searcher should start the search process in the transient state σ_i^t. The only mixed states in the diagram are σ_{-1}^r and σ_{+1}^r, which are entered only after the proper moving state has been occupied. Any probability distribution over their associated alternatives will produce the payoff U(P) as long as the evader uses his good moving strategy.

3-22-4137

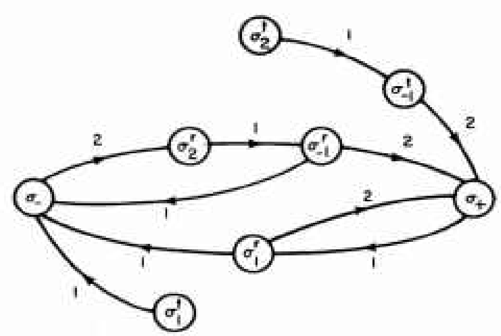
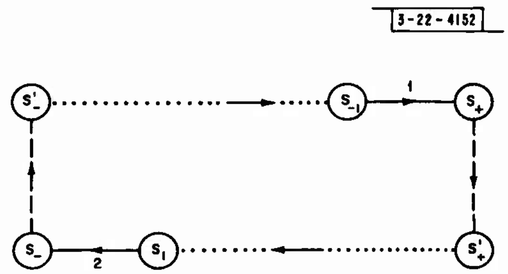


Fig. 16. The searcher's transition diagram.

This transition diagram is typical of those which apply when both P_- and P_+ belong to the interior of the recurrent region (P_{01}, P_{02}). In this situation, both s_- and s_+ are included in the chain diagram associated with game F. In general, such a chain diagram must be of the following form.



Here, of course, s_{-1} may be equivalent to s'_-, and s_{+1} may be equivalent to s'_+. Both always occur, for example, in the first strategy interval. As in the previous example, each state s_i of the no-move region is replaced by two search states σ_i^t and σ_i^r. The recurrent chain is identical to the above chain diagram except that each pair of moving states in the chain diagram is replaced by a single moving state. Each transient state is connected to σ_- or σ_+ in exactly the same manner that s_i is connected to s_- or s_+. In both σ^r_{-1} and σ^r_{+1} a look into either box is admissible. A look into box 1 produces a transition to σ_- and a look into box 2 produces a transition to σ_+.

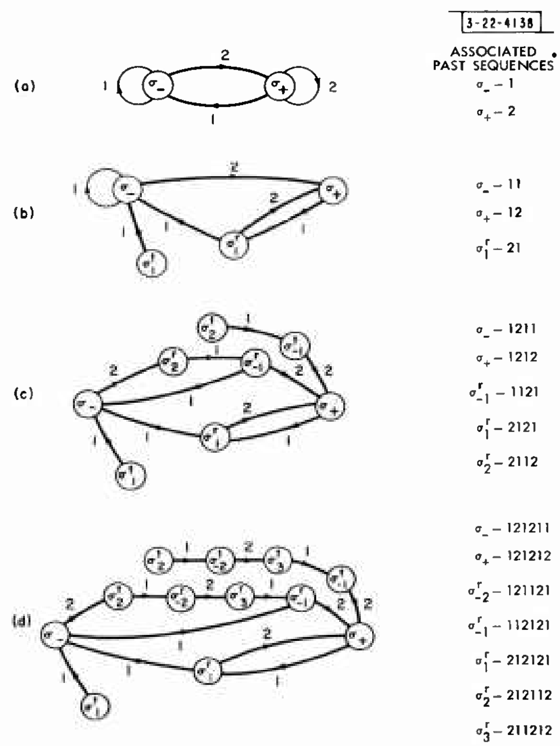
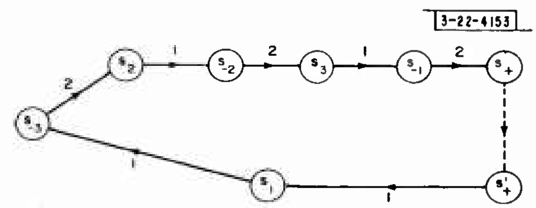


Fig. 17(a–d). The transition diagrams associated with the chain diagrams of Figs. 10(a–d).

The transition diagrams associated with the chain diagrams in Figs. 10(a) through (d) are shown in Fig. 17. Note that each transient chain is used to bypass the mixed states until one of the moving states has been entered.

It is not surprising to find, in each of these diagrams, that a finite part of the past search sequence uniquely determines the recurrent state in which the process must be. This should have been expected, since each state is a behavioral state of the searcher's strategy and must have a corresponding information set. The above property holds for any transition diagram that can be associated with a good search strategy. In Fig. 17(a), each state is defined by the last look; σ_- applies if the last look was made into box 1 and σ_+ applies if it was made into box 2. In Figs. 17(b) through (d), the past sequence of the last two, four, and six looks, respectively, is required to determine uniquely the recurrent state that the process must occupy. Note that not all the possible past sequences of a given number of looks have corresponding states in the diagram. This is true because the pure states cause some sequences to be inadmissible. For example, in each of these latter diagrams it is inadmissible to make two consecutive looks into box 2. Naturally, the finite past sequence associated with each recurrent state is valid only when the search process has generated the required minimum number of looks.

When only one of the bounding points of the no-move region belongs to the interior of the recurrent region, the transition diagram is only slightly different. As an example, consider the chain diagram in Fig. 10(e):



In this case, π_+ extends into the recurrent region and s_+ and s'_+ are included in the chain diagram. Once a move occurs in the transition from s_+ to s'_+, the point P will always be at the left boundary of each interval π_i. As a result, P will be equal to P_0 when it is associated with the interval π_1, and a look into either box will be admissible in σ_1^r. Point P will no longer be equal to P_0 when it belongs to π_{-1} because the moving state s_- does not occur in the chain diagram.

The searcher's transition diagram is shown in Fig. 18. Here, σ_1^t is the only transient state, for all of the other states transform into σ_+ before reaching the single mixed state σ_1^r. This, of course, does not always occur. In general, any other interval π_i that transforms into π_1 before reaching π_{-1} (and hence π_+) requires a transient state σ_i^t in addition to a recurrent state σ_i^r. In

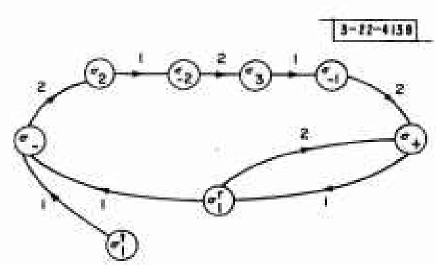


Fig. 18. Transition diagram associated with the chain diagram of Fig. 10(e).

such a situation, σ_i^t will be connected by a series of transitions to σ_1^t in the same manner as σ_i^r is connected to σ_1^r. As before, if P initially belongs to π_i, the search process should start in σ_i^t if such a state exists. If it does not exist the process should start in the unique state σ_i, which is associated with π_i and belongs to the recurrent chain.

5.5 CALCULATION OF THE GOOD PROBABILITY DISTRIBUTIONS ASSOCIATED WITH EACH MIXED STATE

Once the correct transition diagram for the Markov process is determined, the good probability distributions associated with each mixed state must be calculated. We have seen that a payoff equal to U(P) will result for any set of probability distributions as long as the process is started in the correct state and the evader uses his good moving strategy. The good probability distributions that we now seek must limit the evader to this payoff even if he is no longer required to use his good strategy. We must still require him to reveal the initial value of P to the searcher. As would be expected, the modified games H and H' will be of use once we add this constraint.

In order to avoid confusion, let us first consider the case where both P_- and P_+ belong to the interior of the recurrent region and use our standard example, i.e., the chain diagram associated with the modified games F and F' and the transition diagram associated with the modified games H and H' where the initial P is known (Fig. 19).

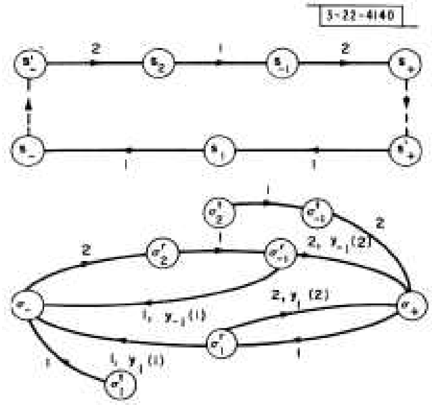


Fig. 19. Chain and transition diagrams associated with the strategy interval of Fig. 15.

As we have seen, the search process should start in σ_i^t if the initial P belongs to π_i, in σ_+ if it belongs to π_+, and in σ_- if it belongs to π_-. Since the evader is no longer required to use his good strategy, it should be clear that the good probability distributions associated with σ_{-1}^r and σ_1^r must insure that $W_i^t(P) = U_i(P)$ for each σ_i^t, that $W_-(P) = U_-(P)$, and that $W_+(P) = U_+(P)$. It should be recalled that $W_i(P)$ and $W_i'(P)$ were defined as the payoffs that result if the evader uses an optimum future strategy. The evader's optimum strategy has more freedom than his good strategy because he knows the strategy used by the searcher. If the searcher does not use his good strategy, the evader can capitalize on this error.

It is easily shown that the payoff $W_i^t(P)$ associated with each transient state is identical to the corresponding payoff $U_i(P)$ as long as $W_-(P) = U_-(P)$ and $W_+(P) = U_+(P)$. If the search process starts in a transient state, only transient states are occupied until σ_- or σ_+ is entered. Until this occurs, all looks are deterministic, and the resulting sequence is the same as that which transforms the equivalent state s_i and s_+ or s_-. Therefore, any $W_i^t(P)$ must be related to $W_-(P)$ or $W_+(P)$ as far as the look sequence is concerned in exactly the same manner as $U_i(P)$ is related to $U_+(P)$ or $U_-(P)$. Each payoff $U_i(P)$ is appropriate to an interval π_i that belongs to the no-move region where $U_i(P) = U_i'(P)$ and $|dU_i'(P)/dr| < \mu$. Therefore, as long as $W_-(P) = U_-(P)$ and $W_+(P) = U_+(P)$, moving cannot be optimum for the evader in any transient search state either, and $W_i^t(P) = W_i'^t(P) = U_i(P) = U_i'(P)$.

In contrast to this behavior, some moves must be admissible in σ_- and σ_+, since the searcher's good strategy must allow the evader to choose any of the admissible alternatives associated with his good strategy at no loss. The state σ_- corresponds to s_- where the evader increases P to P_-. As a result, both moving from box 2 to box 1 and remaining in the same box must be admissible alternatives in σ_-. The function $W_-'(P)$ must, therefore, have a slope equal to $+\mu$ and be identical to $W_-(P)$. Similar reasoning can be used to show that the slope of $W_+'(P)$ must be equal to $-\mu$ and that $W_+'(P) = W_+(P)$.

The necessary and sufficient condition that the searcher's good strategy must satisfy when the initial P is known should now be clear. A pair of probability distributions $Y_{-1} = (y_{-1}(1), y_{-1}(2))$ and $Y_{+1} = (y_{+1}(1), y_{+1}(2))$ must be found that causes $W_-'(P)$ and $W_+'(P)$ to equal $U_-(P)$ and $U_+(P)$, respectively. If this occurs, $W_-(P)$ and $W_+(P)$ will also be equal to the associated payoffs of game F. Such a condition insures that each payoff associated with a transient state will equal the corresponding payoff $U_i(P)$ and that the searcher will be able to limit the evader to U(P) at the beginning of the game.

Before considering the actual computation of the "good" probability distributions, let us show that they exist. The payoff functions U(P) and U'(P) for our example have the general appearance shown in Fig. 15. For the moment, let us assume that $W_-(P) = U_-(P)$ and $W_+(P) = U_+(P)$ and that moving occurs only in σ_- and σ_+. Feedback occurs in the recurrent chain of the transition diagram and these assumptions will be correct if they are not contradicted by this feedback. If we set $y_1(1)$ equal to one, then $\sigma_+ \rightarrow \sigma_-$ in exactly the same manner as $s_+' \rightarrow s_-$. In this case, $W_+'(P) = U_+'(P) \neq U_+(P)$, a contradiction. On the other hand, consider what happens if $y_1(1) = 0$. In this example, π_2 is the interval immediately to the left of P_+. Therefore, $\pi_2 \rightarrow \pi_{-1}$ in exactly the same manner as $\pi_+' \rightarrow \pi_1$. As a result, when $y_1(1) = 0$, $\sigma_+ \rightarrow \sigma_1^r \xrightarrow{2} \sigma_+$ as $s_2 \rightarrow s_{-1} \xrightarrow{2} s_+$. It follows that $W_+'(P) = U_2(P)$. This is again a contradiction. The functions $U_2(P)$, $U_+(P)$ and $U_+'(P)$, however, all intersect at P_+ and have slopes greater than, equal to, and less than $-\mu$, respectively. Therefore, there must exist a $y_1(1)$ where $0 \leqslant y_1(1) \leqslant 1$ for which the slope of $W_+'(P)$ equals $-\mu$. Since $W_+'(P)$ must also intersect the above functions at P_+, it must be identical to $U_+(P)$ when this occurs. The function $W_+'(P)$ is then equal to $W_+(P)$ and no contradiction results. In a similar manner, the slope of $W_-'(P)$ must change from that associated with $U_-(P)$ to that associated with $U_{-1}(P)$ as $y_{-1}(2)$ goes from one to zero. Therefore, there exists a $y_{-1}(2)$ where $0 \leqslant y_{-1}(2) \leqslant 1$ for which $W_-'(P) = U_-(P) = W_-(P)$.

The above argument was developed by assuming that moving could occur only in σ_- and σ_+. Actually, the result is valid as long as the evader can gain no advantage by moving in any recurrent state, that is, as long as $|dW_i'^r(P)/dP| \leqslant \mu$ for each σ_i^r. If this occurs, $W_i'^r(P) = W_i^r(P)$ for each σ_i^r, and the payoffs associated with each state in the transition diagram are consistent. It can be shown that this final requirement is always satisfied by the probability distributions derived in the above manner, and, therefore, that they yield the searcher's good strategy. The proof, however, is somewhat involved. Since it is not particularly illuminating, it has been put in Appendix C.

In order to illustrate the manner in which the good probability distributions can be calculated, let us again refer to the transition diagram in Fig. 19. In the usual manner, we can let

$$W_i^r(P) = W_i'^r(P) = a_i^r P + b_i^r(1 - P) \quad ,$$

$$W_i^t(P) = W_i'^t(P) = a_i P + b_i(1 - P) \quad ,$$

$$W_-(P) = W'_-(P) = a_- P + b_-(1 - P) \quad ,$$

$$W_+(P) = W'_+(P) = a_+ P + b_+(1 - P) \quad .$$

Here, the coefficients associated with the payoffs of the transient states and the two moving states have no superscripts and are identical to those associated with U(P). It follows from Eq. (5-1) that

$$a_1^r = y_1(1)\ (1 + r_1 a_-) + y_1(2)\ (1 + a_+) \quad ,$$

$$b_1^r = y_1(1)\ (1 + b_-) + y_1(2)\ (1 + r_2 b_+) \quad .$$

Since the transition from σ_+ to σ_1^r involves deterministic looks only, Eq. (4-8) can be used to express a_+ in terms of a_1^r and b_+ in terms of b_1^r. In this example, $\sigma_+ \xrightarrow{1} \sigma_1^r$ and we find that

$$a_+ = q_1 + r_1(1 + a_1^r)$$

$$b_+ = 1 + b_1^r \quad .$$

In these equations, the only unknowns are $y_1(1)$, $y_1(2)$, a_1^r and b_1^r. Since $y_1(1) + y_1(2) = 1$, one of the four equations is redundant. As long as the solution of game F is correct, however, no contradiction will result. Since the equations are linear, $Y_1 = (y_1(1),\ y_1(2))$ is easily calculated. The result when $r_1 = 0.512$, $r_2 = 0.4096$ and $\mu = 1.3$ is $Y_1 = (0.4334,\ 0.5666)$.

The quantity Y_2 can be calculated in the same manner by noting that $\sigma_- \xrightarrow{21} \sigma_{-1}^r$. Therefore,

$$a_- = q_1(2) + r_1(2 + a_{-1}^r) \quad ,$$

$$b_- = q_2(1) + r_2(2 + b_{-1}^r) \quad ,$$

while

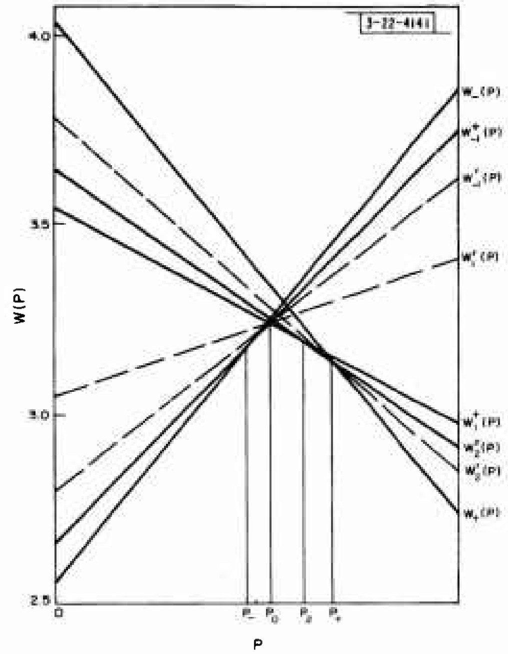
$$a_{-1}^r = y_{-1}(1)\ (1 + r_1 a_-) + y_{-1}(2)\ (1 + a_+) \quad ,$$

$$b_{-1}^r = y_{-1}(1)\ (1 + b_-) + y_{-1}(2)\ (1 + r_2 b_+) \quad .$$

The solution in the numerical example is $Y_{-1} = (0.1575,\ 0.8425)$. It should not come as a surprise to find that $y_{-1}(2) \geqslant y_1(2)$ and that $y_{-1}(1) \leqslant y_1(1)$. This is true in general.

In Fig. 20, the payoff associated with each of the states in the transition diagram is graphed for this numerical example. Those associated with the transient states and the moving states are shown by solid lines, and those associated with the other recurrent states are shown by broken

Fig. 20. The payoff functions associated with the states in the transition diagram shown in Fig. 19 ($r_1 = 0.512$, $r_2 = 0.4096$, $\mu = 1.3$).



lines. Here, of course, each payoff is valid for all P, and there is no need to differentiate between a pair of payoffs for H and H' since they are identical. The lower bound of this ensemble of functions forms the payoff U(P). This follows from the fact that the searcher can limit the evader to U(P) when P initially belongs to π_i only by starting the search process in σ_i^t. Finally, it should be noted that $W_{-1}^t(P) = W_{-1}^r(P)$ and $W_1^t(P) = W_1^r(P)$ at P_0, while $W_2^t(P) = W_2^r(P)$ at P_+. This is true because these are the respective values of P that apply in σ_{-1}^r, σ_1^r and σ_2^r when the evader uses his good strategy.

When only one of the bounding points of the no-move region belongs to the interior of the recurrent region, the searcher's good strategy can be derived in much the same manner. Only one probability distribution is required, however, for there is only one mixed state as well as a single moving state. In the transition diagram of Fig. 18, the payoffs associated with σ_{-3}, σ_2, σ_{-2}, σ_3 and σ_{-1} are each identical to the corresponding $U_i(P)$ in game F, since each of these states is transformed into σ_+ before reaching the mixed state σ_1^r. The function $W_1^r(P)$ will of course be unequal to $U_1(P)$. In general, any other recurrent state σ_i^r that transforms into the mixed state will have an associated transient state and $W_i^r(P) \neq W_i^t(P) = U_i(P)$. Appendix C shows that $|dW_i'^{r}(P)/dP| \leqslant \mu$ for each state of this type also.

5.6 COMPLETION OF THE SEARCHER'S GOOD STRATEGY WHEN INITIAL P IS UNKNOWN

Now that we have seen how the searcher can limit the evader to U(P) when he knows the evader's initial choice of P, we must extend this strategy to the actual game, where initial P is unknown. Clearly, the searcher can no longer limit the evader to U(P) given any P. This is not necessary, however, for we know that the evader can guarantee himself U(P*). As long as we can find a search strategy that limits him to this payoff, U(P*) must be the value of the game and the searcher has a complete good strategy.

All that remains to complete the solution is the computation of the starting rule Y_0 for the Markov process that generates the search sequence. No look is associated with this starting rule, and Y_0 is derived in exactly the same manner as the searcher's starting rule in G^∞. If the evader is to guarantee a payoff equal to $U(P^*)$, he must initially hide with probability P^*. Therefore, if P^* is the breakpoint that separates π_i from π_j, a choice of either σ_i^t or σ_j^t is admissible for the searcher. Since $W_i^t(P)$ and $W_j^t(P)$ must be equal to $U(P^*)$ at P^* and must have slopes of opposite sign, there must exist a $Y_0 = (y_0(\sigma_i^t),\ y_0(\sigma_j^t))$ which insures that $W_0(P) = U(P^*)$ for all P. If, on the other hand, the unusual occurs and U(P) is a maximum over a whole interval π_i, the searcher's starting rule is deterministic and requires the Markov process to start in σ_i^t. Here again the evader is limited to a payoff equal to the maximum of U(P). We may finally state with assurance that a value exists for our search evasion game and that the strategies we have developed for the two players are indeed good strategies.

In the numerical example that has been used throughout these chapters, P^* occurs at $P_0 = 0.538$. Since $W_{-1}^t(P) = 3.7471P + 2.658(1 - P)$ and $W_1^t(P) = 2.9744P + 3.556(1 - P)$, the starting rule requires that $y_0(\sigma_{-1}^t) = 0.3481$ and $y_0(\sigma_1^t) = 0.6519$. The searcher's complete good strategy, illustrated in Fig. 21, limits the evader to $W_0(P) = U(P^*) = 3.243$.

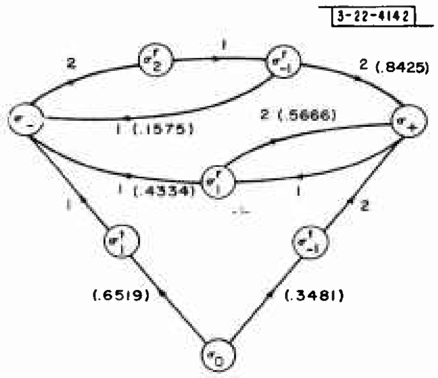


Fig. 21. The searcher's complete good strategy ($r_1 = 0.512$, $r_2 = 0.4096$, $\mu = 1.3$).

CHAPTER 6
GENERALIZED REWARD STRUCTURE

6.1 INTRODUCTION

The two-box search evasion game considered in Chapters 2 through 5 had a very simple reward structure. The evader simply received one unit from the searcher each time a look was made and paid him μ units each time he moved. Thus, the reward associated with each look was independent of where the look was made and where the evader was hiding at the time. Also, the moving cost was not a function of the direction of the move.

We have deferred treating the more general reward structure until now because it has allowed us to study the behavior of the game with a simpler notation. In this chapter, we shall examine the two-box game with a more general reward structure. Most of the properties that have been developed will carry over directly. In fact, all the properties that make the two-box search evasion game interesting have already appeared. These properties arose because

(a) The searcher did not know where the evader was until he found him.

(b) The state variable P was changed according to Bayes' rule by each unsuccessful look and this transformation was a function of the escape probabilities alone.

(c) The evader could move at a cost, and the cost of a transformation of the state variable P was proportional to the magnitude of the change in P.

These properties will still apply.

In the example of revenuer vs moonshiner in Chapter 1, we noted that a reward of one unit was associated with each look. It took the revenuer one time unit to examine an area, and during this time the moonshiner was able to produce enough moonshine to secure one unit of profit. We can imagine that in a more general situation it takes the revenuer different amounts of time to examine the various areas. Also, the moonshiner may be able to operate more efficiently in one area than in another. That is, his earning rate may vary from box to box. As a result, the reward associated with a given look may depend on where the look is made and where the moonshiner, or evader, is hiding.

To account for these possibilities, as well as others, let us introduce the following reward structure:

ρ_i = evader's earning rate in box i if the searcher is not looking there;

η_i = loss in earning rate in box i when the searcher is looking there (net earning rate = $\rho_i - \eta_i$);

τ_i = time required to examine box i;

λ_i = detection loss of box i.

In order to make these quantities realistic, we shall require that $\rho_i, \tau_i > 0$; $\eta_i, \lambda_i \geqslant 0$.

Our reward structure can be interpreted as follows. In the event that the searcher looks into box j while the evader is hiding in box i, the evader receives the reward $\rho_i \tau_j$. If, on the other hand, the searcher looks into box i, the evader receives a reward of $(\rho_i - \eta_i)\,\tau_i$. The introduction of η_i allows us to consider examples in which the evader cannot operate as efficiently when the searcher is examining the box in which he is hiding. To be realistic, η_i should not exceed ρ_i; that is, $\rho_i - \eta_i \geqslant 0$. However, this restriction will not be formally imposed for it is not necessary in the mathematical development.

The net reward associated with the event in which the searcher looks into box i and finds the evader need not be equal to $(\rho_i - \eta_i)\,\tau_i$. The evader may be found at the beginning of this look. In addition, he may suffer a penalty for being caught. For example, the evader may be sent to jail. We can combine these losses in λ_i, the detection loss. The net reward associated with the event in which the searcher looks into box i and finds the evader is, therefore, $(\rho_i - \eta_i)\,\tau_i - \lambda_i$.

A final generalization that can be applied to the reward structure with little increase in complexity concerns the moving cost. We can let the moving cost depend on which move is made. Since only two boxes are considered here, only two moves can occur. Thus we can let μ_1 represent the cost associated with a move to box 1 from box 2 and μ_2 represent the cost of the move in the reverse direction. The subscripts of these coefficients correspond to those associated with the move probabilities x_1 and x_2.

In order to utilize the work of the previous chapters most efficiently, the two-box search evasion game with the generalized reward structure will be discussed in the same sequence. Those properties that still hold will be mentioned, exceptions will be noted, and the new form of each of the various equations that were of use before will be listed. To simplify the association of each new form with the old, each new equation will be numbered as before but will be followed by the symbol §.

6.2 G^∞: THE NO-MOVE GAME

When moving is not allowed, G^∞ can be solved as before by using the modified game F^∞. The function $U^\infty(P)$ is continuous and convex. It is piecewise linear under the conditions stated in Chapter 2. A single infinite search sequence is optimum over any interval in P over which $U^\infty(P)$ is linear. The fundamental recursion equation that applies in place of Eq. (2-4) is

$$U^\infty(P) = \min \begin{cases} U^\infty(P;1) = P[(\rho_1 - \eta_1)\,\tau_1 - q_1\lambda_1] + (1-P)\,\rho_2\tau_1 \\ \qquad + [Pr_1 + 1 - P]\,U^\infty\left[\dfrac{Pr_1}{Pr_1 + 1 - P}\right] \\ U^\infty(P;2) = P\rho_1\tau_2 + (1-P)[(\rho_2 - \eta_2)\,\tau_2 - q_2\lambda_2] \\ \qquad + [P + (1-P)\,r_2]\,U^\infty\left[\dfrac{P}{P + (1-P)\,r_2}\right] \end{cases} \quad . \qquad (2\text{-}4)§$$

The searcher's optimum strategy requires that

if $P > P_0$, look into box 1 ;

if $P < P_0$, look into box 2 ;

if $P = P_0$, look into either box .

Again, this property is derived rigorously in Appendix A. The point P_0 can be calculated again by requiring that $U^\infty(P_0;12) = U^\infty(P_0;21)$:

$$U^{\infty}(P;12) = \{P[(\rho_1 - \eta_1)\ \tau_1 - q_1\lambda_1] + (1-P)\ \rho_2\tau_1\}$$

$$+ \{Pr_1\rho_1\tau_2 + (1-P)\ [(\rho_2 - \eta_2)\ \tau_2 - q_2\lambda_2]\}$$

$$+ [Pr_1 + (1-P)\ r_2]\ U^{\infty}\left[\frac{Pr_1}{Pr_1 + (1-P)\ r_2}\right] \ ,$$

$$= U^{\infty}(P;21) = \{P\rho_1\tau_2 + (1-P)\ [(\rho_2 - \eta_2)\ \tau_2 - q_2\lambda_2]\}$$

$$+ \{P[(\rho_1 - \eta_1)\ \tau_1 - q_1\lambda_1] + (1-P)\ r_2\rho_2\tau_1\}$$

$$+ [Pr_1 + (1-P)\ r_2]\ U^{\infty}\left[\frac{Pr_1}{Pr_1 + (1-P)\ r_2}\right]$$

implies that

$$P_0 = \frac{\dfrac{\tau_1}{\rho_1 q_1}}{\dfrac{\tau_1}{\rho_1 q_1} + \dfrac{\tau_2}{\rho_2 q_2}} \ .$$

It is interesting to note that P_0 is independent of η_i and λ_i. Since ρ_i and τ_i must be positive but finite, P_0 will always lie in the interior of the interval (0, 1).

The transformation of the state variable P is a function of r_1 and r_2 (or q_1 and q_2) only; therefore the recurrent region (P_{01}, P_{02}) can be defined as in Chapter 2. Once P enters this region it must remain in it as long as the searcher uses an optimum strategy. It is possible to calculate $U^{\infty}(P)$ outside of this region once $U^{\infty}(P)$ is known within it. The payoff inside the recurrent region can be calculated in the same manner as before because the optimum chain diagram remains the same. Only the position of P_0 and, therefore, the other breakpoints are functions of ρ_i and τ_i.

When a chain diagram is used to generate the search sequence, a linear payoff $U_i^{\infty}(P) = a_i P + b_i(1-P)$ can again be associated with each state s_i in the chain. If the chain is optimum, $U_i^{\infty}(P)$ will be equal to the optimum payoff over the associated interval π_i.

The equations which relate the payoff associated with one state to that of another are

$$s_i \xrightarrow{1} s_j \Longrightarrow$$

$$a_i = (\rho_1 - \eta_1)\ \tau_1 - q_1\lambda_1 + r_1 a_j \ ,$$

$$b_i = \rho_2\tau_1 + b_j \ ;$$

$$s_i \xrightarrow{2} s_j \Longrightarrow$$

$$a_i = \rho_1\tau_2 + a_j \ ,$$

$$b_i = (\rho_2 - \eta_2)\ \tau_2 - q_2\lambda_2 + r_2 b_j \ . \qquad (2\text{-}11)\S$$

When s_i is transformed into s_j by a sequence defined by $\{t_m(n)\}$ which involves a total of k_1 looks into box 1 and k_2 looks into box 2, we have

$$s_i \xrightarrow{\{t_m(n)\}} s_j \Longrightarrow$$

$$a_i = q_1 \sum_{n=1}^{k_1} r_1^{n-1} \{(\rho_1 - \eta_1)\, n\tau_1 + \rho_1 [t_1(n) - n]\, \tau_2 - \lambda_2\}$$

$$+ r_1^{k_1} [(\rho_1 - \eta_1)\, k_1 \tau_1 + \rho_1 k_2 \tau_2 + a_j] \quad ,$$

$$b_i = q_2 \sum_{n=1}^{k_2} r_2^{n-1} \{(\rho_2 - \eta_2)\, n\tau_2 + \rho_2 [t_2(n) - n]\, \tau_1 - \lambda_2\}$$

$$+ r_2^{k_2} [\rho_2 k_1 \tau_1 + (\rho_2 - \eta_2)\, k_2 \tau_2 + b_j] \quad . \qquad (2\text{-}13)\S$$

By letting $\{t_m(n)\}$ represent an infinite search sequence, it becomes clear that the payoff associated with such an infinite sequence must be linear in P. Finally, if $r_1^{n_1} = r_2^{n_2}$, the optimum search sequence will be periodic inside the current region and we find that

$$s_i \xrightarrow{\{t_m(n)\}} s_i \Longrightarrow$$

$$a_i = \frac{1}{1 - r_1^{n_1}} \left(q_1 \sum_{j=1}^{n_1} r_1^{j-1} \{(\rho_1 - \eta_1)\, j\tau_1 + \rho_1 [t_1(j) - j]\, \tau_2 - \lambda_1\} \right.$$

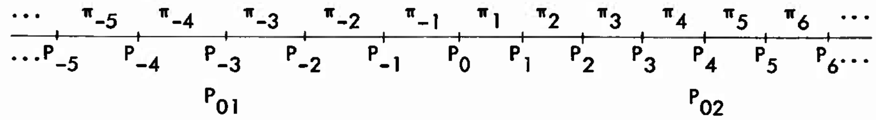
$$\left. + r_1^{n_1} [(\rho_1 - \eta_1)\, n_1 \tau_1 + \rho_1 n_2 \tau_2] \right) ,$$

$$b_i = \frac{1}{1 - r_2^{n_2}} \left(q_2 \sum_{j=1}^{n_2} r_2^{j-1} \{(\rho_2 - \eta_2)\, j\tau_2 + \rho_2 [t_2(j) - j]\, \tau_1 - \lambda_2\} \right.$$

$$\left. + r_2^{n_2} [\rho_2 n_1 \tau_1 + (\rho_2 - \eta_2)\, n_2 \tau_2] \right) .$$

The only important differences which arise with the introduction of the generalized reward structure are that $U^\infty(P)$ may be negative for some P and that it need not achieve its maximum inside the recurrent region. The former situation can occur if one or both of the detection losses is very large. This has no other effect on the solution, although it may deter the evader from playing the game. Furthermore, no difficulty should be encountered if $U^\infty(P)$ is a maximum outside the recurrent region. The payoff inside the recurrent region can be calculated in exactly the same manner as before, and once this has been done, it can be calculated as far into a transient region as is necessary.

To illustrate the manner in which the solution can be extended into the transient regions, let us consider our familiar example where $r_1^4 = r_2^3$. The interval (0, 1) can be partitioned into the intervals over which $U^\infty(P)$ is linear as before:



The same chain diagram also applies for intervals inside the recurrent region. Transient states can be added by noting that a look into box 1 shifts an interval n_2 places to the left and a look into box 2 shifts an interval n_1 places to the right. The chain diagram that includes some of the transient states is shown in Fig. 22. Once the payoff associated with each state in the recurrent chain

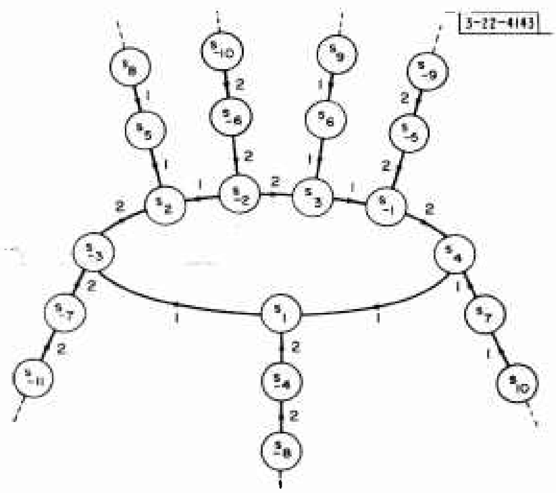


Fig. 22. A chain diagram with transient states.

is known, those associated with each transient state can be calculated by using Eq. (2-11)§ or Eq. (2-13)§. The values of the separating breakpoints can be calculated in exactly the same way as before. Naturally, the solution should be extended only in the direction in which $U^\infty(P)$ increases and only as far as is necessary.

The point P^* is again the evader's good strategy in G^∞. Once the payoffs associated with the two states that are optimum at $P = P^*$ have been found, the searcher's good strategy can be calculated. Equation (2-15) can be used to make this computation without alteration.

6.3 GAME G°: $\mu_1, \mu_2 = 0$

When both moving costs are equal to zero, the game can be solved as it was in Chapter 3. The evader should restore the state variable P to its optimum value P_0 after each unsuccessful look, and the searcher should make each look according to the probability distribution $\underline{Y}_0 = (y_0(1), y_0(2))$.

If the evader always restores the state variable to P before each next look and the searcher looks into box 1,

$$U^\circ(P; 1) = P[(\rho_1 - \eta_1)\,\tau_1 - q_1\lambda_1] + (1 - P)\,\rho_2\tau_1 + (Pr_1 + 1 - P)\,U^\circ(P) \quad .$$

If a look into box 1 is optimum for a given P, it is always optimum for that P; therefore,

$$U^\circ(P;1) = \frac{1}{Pq_1}\left\{P[(\rho_1 - \eta_1)\tau_1 - q_1\lambda_1] + (1-P)\rho_2\tau_1\right\} \quad .$$

Similarly, if the evader always returns the state variable to P before each look and the searcher always looks into box 2, the payoff is

$$U^\circ(P;2) = \frac{1}{(1-P)q_2}\left\{P\rho_1\tau_2 + (1-P)[(\rho_2 - \eta_2)\tau_2 - q_2\lambda_2]\right\} \quad .$$

The optimum value P_0 of the state variable is that which maximizes the minimum of $U^\circ(P;1)$ and $U^\circ(P;2)$. Since both of these functions are nonlinear, it must be shown that there indeed exists a P_0 where $0 < P_0 < 1$ for which $\max_P (\min_i U(P;i)) = U^\circ(P_0;1) = U^\circ(P_0;2)$. The demonstration is carried out in Appendix D. It follows that the evader's good strategy and the resulting guaranteed payoff can be found by solving the equations

$$U^\circ(P_0) = \frac{1}{P_0q_1}\left\{P_0[(\rho_1 - \eta_1)\tau_1 - q_1\lambda_1] + (1-P_0)\rho_2\tau_1\right\}$$

$$= \frac{1}{(1-P_0)q_2}\left\{P_0\rho_1\tau_2 + (1-P_0)[(\rho_2 - \eta_2)\tau_2 - q_2\lambda_2]\right\} \quad .$$

The nonlinearity of the functions $U^\circ(P;1)$ and $U^\circ(P;2)$ might appear surprising at first thought, because similar functions such as $U'(P;i)$ in Chapter 4 have usually been linear or piecewise linear. Note, however, that in that chapter $U'(P;i)$ was defined as the payoff in F' which resulted if the searcher looked first into box i and both players used optimum strategies thereafter. The function $U^\circ(P;i)$, on the other hand, has been defined here as the payoff that results if the searcher always looks into box i and the evader always returns the state variable to P. Thus, the evader's entire future strategy is a function of the variable P. When P is unequal to P_0, the evader's entire future strategy is not optimum. This point was not mentioned in Chapter 2 since the general game with $\mu \neq 0$ had not been considered at that time.

Let us return to G°, where the searcher's good strategy can be found in the same manner as was the evader's. If the searcher uses the probability distribution $\underline{Y} = (Y, 1 - Y)$ and the evader hides in box 1, the payoff is

$$W^\circ(Y;1) = Y[(\rho_1 - \eta_1)\tau_1 - q_1\lambda_1] + (1-Y)\rho_1\tau_2 + [Yr_1 + (1-Y)]W^\circ(Y) \quad .$$

If, on the other hand, the evader hides in box 2, we find that

$$W^\circ(Y;2) = Y\rho_2\tau_1 + (1-Y)[(\rho_2 - \eta_2)\tau_2 - q_2\lambda_2] + [Y + (1-Y)r_2]W^\circ(Y) \quad .$$

The searcher's good strategy can be found by solving the equation

$$W^\circ(Y_0) = \frac{1}{Y_0q_1}\left\{Y_0[(\rho_1 - \eta_1)\tau_1 - q_1\lambda_1] + (1-Y_0)\rho_1\tau_2\right\}$$

$$+ \frac{1}{(1-Y_0)q_2}\left\{Y_0\rho_2\tau_1 + (1-Y_0)[(\rho_2 - \eta_2)\tau_2 - q_2\lambda_2]\right\} \quad ,$$

because the above equations yield a Y_0 for which $0 < Y_0 < 1$ and $W(Y_0) = U(P_0) = V^\circ$, the value. These properties are also shown in Appendix D.

Since the equations used to determine P_0, Y_0 and $V°$ are nonlinear, it is not convenient to express the solution algebraically. Nonetheless, the solutions are rather easy to obtain numerically.

In several special cases algebraic solutions can be found readily. For example, if $\eta_i = \lambda_i = 0$,

$$P_0 = \frac{\tau_1/q_1}{\tau_1/q_1 + \tau_2/q_2} \quad , \quad Y_0 = \frac{\rho_1/q_1}{\rho_1/q_1 + \rho_2/q_2} \quad , \quad V° = \frac{\tau_1\rho_1}{q_1} + \frac{\tau_2\rho_2}{q_2} \quad .$$

When $\lambda_i = 0$ and $\eta_i = \rho_i$ [when $(\rho_i - \eta_i)\,\tau_i = 0$], then

$$P_0 = \frac{\sqrt{\frac{\tau_1}{\rho_1 q_1}}}{\sqrt{\frac{\tau_1}{\rho_1 q_1}} + \sqrt{\frac{\tau_2}{\rho_2 q_2}}} \quad , \quad Y_0 = \frac{\sqrt{\frac{\rho_1}{\tau_1 q_1}}}{\sqrt{\frac{\rho_1}{\tau_1 q_1}} + \sqrt{\frac{\rho_2}{\tau_2 q_2}}} \quad , \quad V° = \sqrt{\frac{\tau_1\tau_2\rho_1\rho_2}{q_1 q_2}} \quad .$$

In this last example, the searcher concentrates more attention on a box if the earning rate is large and the look time is small. The evader does the reverse. Both players, however, concentrate more attention on a box if its detection probability is small.

6.4 GAME G

When the evader can move between looks at a cost, the search evasion game can be solved in essentially the same manner as in Chapters 4 and 5. The efficient move condition given in Eq. (4-1) still holds, but now the cost of the transformation of the state variable depends on whether it is increased or decreased; that is,

$$C(P \to P') = \begin{cases} \mu_1(P' - P) & , \quad P' \geqslant P \\ \mu_2(P - P') & , \quad P' \leqslant P \quad . \end{cases} \tag{4-2)§$$

As was mentioned, μ_1 is the cost associated with a move to box 1 from box 2 and μ_2 applies for a move in the reverse direction.

The modified games F and F' can be used as before, but the functional equations are now

$$U'(P) = \min \begin{cases} U'(P;1) = P[(\rho_1 - \eta_1)\,\tau_1 - q_1\lambda_1] + (1 - P)\,\rho_2\tau_1 \\ \qquad + [Pr_1 + 1 - P]\,U\left[\dfrac{Pr_1}{Pr_1 + 1 - P}\right] \\ U'(P;2) = P\rho_1\tau_2 + (1 - P)\,[(\rho_2 - \eta_2)\,\tau_2 - q_2\lambda_2] \\ \qquad + [P + (1 - P)\,r_2]\,U\left[\dfrac{P}{P + (1 - P)\,r_2}\right] \end{cases} \tag{4-3)§$$

and

$$U(P) = \max_{P'} \begin{cases} -\mu_1(P' - P) + U'(P') & , \quad P' \geqslant P \\ -\mu_2(P - P') + U'(P') & , \quad P' \leqslant P \quad . \end{cases} \tag{4-4)§$$

Both U(P) and U'(P) must still be continuous and convex (see Appendix B). As in Fig. 5, U(P) and U'(P) are identical over the no-move region (P_-, P_+), but now this region is defined by

$$\frac{dU'(P)}{dP} \geqslant \mu_1 \quad , \quad P \leqslant P_-$$

$$< \mu_1 \quad , \quad P > P_- \quad ;$$

$$-\frac{dU'(P)}{dP} < \mu_2 \quad , \quad P < P_+$$

$$\geqslant \mu_2 \quad , \quad P \geqslant P_+ \quad . \qquad (4\text{-}5)\S$$

The function U(P) has a slope equal to $+\mu_1$ in the moving region $(0, P_-)$ and a slope of $-\mu_2$ in the moving region $(P_+, 1)$.

The fundamental property of the searcher's optimum strategy in F', which was discussed in Sec. 4.3, remains the same (see Appendix B). There exists a P_0, where $P_- \leqslant P_0 \leqslant P_+$, such that,

$$P < P_0 \Longrightarrow U'(P;2) < U'(P;1) \Longrightarrow \text{look into box 2} \quad ,$$

$$P > P_0 \Longrightarrow U'(P;1) < U'(P;2) \Longrightarrow \text{look into box 1} \quad .$$

The moving costs are again prohibitive and the game can be solved in terms of G^∞ if the no-move region contains the recurrent region. This occurs if

$$\mu_1 \geqslant \left.\frac{dU^\infty(P)}{dP}\right|_{P=P_{01}^+}$$

$$\mu_2 \geqslant -\left.\frac{dU^\infty(P)}{dP}\right|_{P=P_{02}^-} \quad . \qquad (4\text{-}7)\S$$

When this condition holds, U(P) will be identical to $U^\infty(P)$ over the no-move and hence the recurrent region. Although U(P) may be a maximum outside the recurrent region, clearly it must achieve this maximum value inside the no-move region where it is identical to $U^\infty(P)$. Thus, the value and good strategies can be obtained from the function $U^\infty(P)$ as before. It should be noted that a simple prohibitive bound cannot be placed on either moving cost. The moving costs can be considered prohibitive in the previous sense only if they both satisfy (4-7)§.

When the moving costs are not prohibitive, the correct chain diagram must be found before the payoff function U(P) can be calculated. This can again be accomplished by studying the manner in which the form of the payoff functions changes from strategy interval to strategy interval as μ_1 and μ_2 increase up to their appropriate values. In order to do this, it is best to hold μ_1 and μ_2 in a fixed ratio as they are increased.

Although the form of the chain diagram associated with the optimum search strategy is independent of the reward coefficients when the moving costs are prohibitive, this is not true when the moving costs are not prohibitive. In Chapter 4, we saw that two possible changes could occur in the form of the payoff function U(P), and hence in the form of the chain diagram, at the end of each strategy interval. The change that now occurs depends on the reward coefficients and the ratio μ_1/μ_2 that are used. Thus, the sequence of chain diagrams associated with a particular pair of detection probabilities and the simple reward structure need not be the same as that which occurs when an arbitrary set of reward coefficients is used.

When both moving regions extend into the recurrent region and the appropriate chain diagram is known, the payoff functions can be calculated in much the same way as in Sec. 4.7. The chain diagram must include the moving states s_- and s_+ and each interval in the no-move region must have an associated state in the chain. As a result, the two search states optimum at P* must be included in the chain and no additional transient states are required. Once the payoff coefficients have been introduced, the same set of equations may be used to obtain a solution. The function $U'(P)$ may be expressed in terms of $U_+(P)$ and $U'_+(P)$ may be expressed in terms of $U_-(P)$ by means of Eq. (4-8)§. Equation (4-8)§ is identical to Eq. (2-13)§ in Sec. 6.2, just as Eq. (4-8) is identical to Eq. (2-13). Equation (4-9) must be modified to

$$a_- - b_- = \mu_1 \quad ,$$

$$b_+ - a_+ = \mu_2 \quad . \tag{4-9}§$$

The functions $U_-(P)$ and $U'_-(P)$ must again intersect at P_-, and $U_+(P)$ and $U'_+(P)$ must intersect at P_+. Therefore, as previously,

$$a_-P_- + b_-(1 - P_-) = a'_-P_- + b'_-(1 - P_-) \quad ,$$

$$a_+P_+ + b_+(1 - P_+) = a'_+P_+ + b'_+(1 - P_+) \quad . \tag{4-10}$$

The sequences of looks which transform P_- into P_0 and P_+ into P_0 can also be found from the chain diagram, and Eq. (4-11) remains the same as before:

$$P_i \xrightarrow{(k_1, k_2)} P_j \Longrightarrow P_i = \frac{P_j r_2^{k_2}}{P_j r_2^{k_2} + (1 - P_j)\, r_1^{k_1}} \quad . \tag{4-11}$$

Finally, Eq. (4-12) must be modified and we find that

$$P_0[(\rho_1 - \eta_1)\,\tau_1 - q_1\lambda_1 + r_1a_-] + (1 - P_0)\,(\rho_2\tau_1 + b_-) = P_0(\rho_1\tau_2 + a_+)$$
$$+ (1 - P_0)\,[(\rho_2 - \eta_2)\,\tau_2 - q_2\lambda_2 + r_2b_+] \quad . \tag{4-12}§$$

Once this set of equations has been solved, the payoffs associated with the other states in the chain can be calculated by means of Eq. (4-8)§ and the remaining breakpoints can be found by using Eq. (4-11).

If only P_- or P_+ belongs to the interior of the recurrent region, the payoff can be calculated inside the recurrent region in the same manner as in Chapter 4 once the appropriate equations have been modified as above. Although U(P) must always be a maximum at a point that lies in the no-move region, this point need not lie in the recurrent region. Transient states associated with intervals that lie in the no-move but not the recurrent region can be attached to the recurrent chain in exactly the manner discussed in Sec. 6.2. Once this has been done, the payoffs associated with each of these states can be calculated by using Eq. (4-8)§, since no moving occurs during the look sequence that transforms such a state into the recurrent chain.

It is perhaps worth noting here that, with the simple reward structure, intervals may also exist which lie in the no-move but not in the recurrent region. For example, in Fig. 9(e), the bounding point P_- may shift to the left of $P_{01} = P_{-3}$ before P_+ shifts from P_3 to P_4. This situation would have no effect on the recurrent chain in Fig. 10(e). Therefore, it would have no effect on the manner in which the payoffs associated with the states in this chain are calculated. If such a shift were to occur, a linear interval π_{-4} would result. This interval would lie in the no-move but not the recurrent region. The associated state s_{-4} would be a transient state and would be transformed into s_1 by an optimum look into box 2.

The possibility of such behavior was not mentioned in Chapter 4 because it was tacitly assumed that P* would always lie inside the recurrent as well as the no-move region. Although this assumption has not been proved, it is the author's opinion that it is indeed valid. However, if this faith were contradicted by some special example, the result would not be catastrophic, for the payoff associated with a state such as s_{-4} in the above example could be calculated easily. Once this had been done, the evader's good strategy could be calculated as before. The searcher's good strategy also could be found in the usual manner once a transient state σ^t_{-4} was attached to σ^t_1 by a look into box 2 (see Fig. 18). Note that in such a situation there would be no corresponding recurrent state σ^r_{-4}, since only states having associated intervals in the recurrent region belong to the recurrent chain.

In the search evasion game with the generalized reward structure, the searcher's good strategy can be derived easily once games F and F' have been solved. The searcher's good strategy must again be Markovian and the recurrent chain of the transition diagram is identical to the recurrent chain of the chain diagram of games F and F' after the move transitions have been deleted. A transient state σ^t_i is associated with each interval π_i in the no-move region and transforms into σ_- or σ_+ in exactly the same manner as π_i transforms into π_- or π_+. If P_- and P_+ do not both belong to the interior of the recurrent region, there may exist intervals that lie in the no-move but not the recurrent region. As was just mentioned, such an interval will have an associated transient state σ^t_i but not a recurrent state σ^r_i in the transition diagram.

The searcher can limit the evader to U(P) when the initial P is known as long as $U_i(P) = W_i(P)$ for each of the moving states that belongs to the recurrent chain of the transition diagram.

The fundamental functional equations of games H and H' necessary now are

$$W'_i(P) = y_i(1)\left\{\begin{aligned}&P[(\rho_1-\eta_1)\tau_1 - q_1\lambda_1] + (1-P)\rho_2\tau_1\\&+[Pr_1+1-P]\,W_{i|1}\left[\frac{Pr_1}{Pr_1+1-P}\right]\end{aligned}\right\}$$
$$+y_i(2)\left\{\begin{aligned}&P\rho_1\tau_2 + (1-P)[(\rho_2-\eta_2)\tau_2 - q_2\lambda_2]\\&+[P+(1-P)r_2]\,W_{i|2}\left[\frac{P}{P+(1-P)r_2}\right]\end{aligned}\right\}, \qquad (5\text{-}1)§$$

and

$$W_i(P) = \begin{cases} -\mu_2 P + W'_i(0) , & \dfrac{dW'_i(P)}{dP} < -\mu_2 \\ W'_i(P) , & -\mu_2 \leqslant \dfrac{dW'_i(P)}{dP} \leqslant \mu_1 \\ -\mu_1(1-P) + W'_i(1) , & \dfrac{dW'_i(P)}{dP} > \mu_1 \end{cases} \qquad (5\text{-}2)§$$

When both σ_- and σ_+ belong to the recurrent chain of the searcher's transition diagram, the good probability distribution $\underline{Y}_1 = \{y_1(1), y_1(2)\}$ for the mixed state σ_1 can be found from the equations

$$a_1^r = y_1(1)\,[(\rho_1 - \eta_1)\,\tau_1 - q_1\lambda_1 + r_1 a_-] + y_1(2)\,(\rho_1\tau_2 + a_+) \quad ,$$

$$b_1^r = y_1(1)\,(\rho_2\tau_1 + b_-) + y_1(2)\,[(\rho_2 - \eta_2)\,\tau_2 - q_2\lambda_2 + r_2 b_+] \quad ,$$

$$a_+ = q_1 \sum_{n=1}^{k_1} r_1^{n-1}\{(\rho_1 - \eta_1)\,n\tau_1 + \rho_1[t_1(n) - n]\,\tau_2 - \lambda_1\}$$

$$+ r_1^{k_1}[(\rho_1 - \eta_1)\,k_1\tau_1 + \rho_1 k_2\tau_2 + a_1^r] \quad ,$$

$$b_+ = q_2 \sum_{n=1}^{k_2} r_2^{n-1}\{(\rho_2 - \eta_2)\,n\tau_2 + \rho_2[t_2(n) - n]\,\tau_1 - \lambda_2\}$$

$$+ r_2^{k_2}[\rho_2 k_1\tau_1 + (\rho_2 - \eta_2)\,k_2\tau_2 + b_1^r] \quad ,$$

where $\{t_m(n)\}$ represents the sequence in the recurrent chain that transforms σ_+ into σ_1^r and a_+, b_+, a_- and b_- are again the coefficients associated with $U_+(P)$ and $U_-(P)$. In the same manner, $\underline{Y}_{-1} = \{y_{-1}(1), y_{-1}(2)\}$ must satisfy the equations

$$a_{-1}^r = y_{-1}(1)\,[(\rho_1 - \eta_1)\,\tau_1 - q_1\lambda_1 + r_1 a_-] + y_{-1}(2)\,(\rho_1\tau_2 + a_+) \quad ,$$

$$b_{-1}^r = y_{-1}(1)\,(\rho_2\tau_1 + b_-) + y_{-1}(2)\,[(\rho_2 - \eta_2)\,\tau_2 - q_2\lambda_2 + r_2 b_+] \quad ,$$

$$a_- = q_1 \sum_{n=1}^{k_1} r_1^{n-1}\{(\rho_1 - \eta_1)\,n\tau_1 + \rho_1[t_1(n) - n]\,\tau_2 - \lambda_1\}$$

$$+ r_1^{k_1}[(\rho_1 - \eta_1)\,k_1\tau_1 + \rho_1 k_2\tau_2 + a_{-1}^r] \quad ,$$

$$b_- = q_2 \sum_{n=1}^{k_2} r_2^{n-1}\{(\rho_2 - \eta_2)\,n\tau_2 + \rho_2[t_2(n) - n]\,\tau_1 - \lambda_2\}$$

$$+ r_2^{k_2}[\rho_2 k_1\tau_1 + (\rho_2 - \eta_2)\,k_2\tau_2 + b_{-1}^r] \quad ,$$

where $\{t_m(n)\}$ represents the sequence that transforms σ_- into σ_{-1}^r.

If only one of the bounding points of the no-move region belongs to the interior of the recurrent region, only one moving state and one mixed state will occur in the recurrent chain of the transition diagram. The good probability distribution associated with the mixed state can be calculated as in Chapter 5 once the appropriate equations have been modified as they were above.

When the good probability distributions associated with the mixed states have been found, the starting rule $\underline{Y}_0$ needed in G where the initial P is unknown can be calculated exactly as before. The starting rule merely provides a probability distribution for selecting the state in which the Markov process starts. Since a look is not associated with this selection, the equation used to calculate $\underline{Y}_0$ is identical to that used in Chapter 5.

CHAPTER 7
DISCOUNTING

7.1 INTRODUCTION

In the last chapter, the reward structure of the search evasion game was generalized. Nevertheless, the contribution to the payoff of any particular event, i.e., the utility of the event, was still considered independent of when the event occurred. In many cases, however, this is not appropriate. For example, if a reward of one dollar is associated with a given event, the utility of the event should be greater if the event occurs immediately rather than in the future. A dollar in hand can be invested and earn interest.

In this chapter, we shall consider the behavior of the search evasion game when future rewards must be discounted. The term discounting is used when the utility of an event can be found by multiplying the associated reward (the utility that applies when the event occurs) by a discount factor. This discount factor, which is applied to all rewards, must be a function of only the difference between the time at which the utility is evaluated and the time when the event occurs. Thus, it must have the property of stationarity. A further restriction which will be imposed in this chapter is that the discount factor must decay exponentially with time.

Such a discount factor is clearly appropriate when the various rewards are made in monetary units. In the example of revenuer vs moonshiner, this is the case. If the moonshiner is able to invest his profits so that they earn compound interest at a rate α per unit time, one dollar invested at $t = 0$ increases in value according to the function $e^{\alpha t}$. By reversing this reasoning, we find that a reward of one unit received at time t should have a utility at $t = 0$ of $e^{-\alpha t}$. Thus, $e^{-\alpha t}$ is the discount factor. If interest is compounded only at discrete time intervals, as for example in a savings bank, the discount factor does not decay continuously but at discrete intervals. As long as the interest per period is of the order of a few percent or less, the approximation of continuous compounding is very good.

When rewards are not made in monetary units, exponential discounting is often still appropriate and in many other situations it serves as a useful approximation. One must, of course, be careful that the utility of a reward decays in a manner which depends only on the total decay time and not on the time when the decay begins.

As an example of a reward that does not satisfy this requirement of stationarity, consider the utility of information concerning the fixing of a horse race. Such information clearly has a high utility to a prospective wagerer if it is received before the race is run. Once the race is over, however, it has no value at all (except, of course, to a race official). Thus, the utility of such information does not depend upon how far in the future it is received but on when it is received relative to the time of the race.

As we can see, the restriction of stationarity implied by the term discounting is a very strong one. On the other hand, the further requirement that the discount factor decay exponentially with time does not restrict its applicability appreciably more. A little thought will show that in most cases if the utility of a reward decays in a nonexponential manner, the requirement of stationarity itself is actually violated.

As an example of a situation in which exponential discounting of rewards may be appropriate when the rewards do not involve money, consider a search evasion game in which the evader is a clandestine manufacturer of ballistic missiles violating an arms control agreement and in which the searcher is a member of an inspectorate set up to police this agreement. Here, the evader manufactures weapons rather than moonshine. The utility of a given stockpile of missiles must be defined in terms of the political power (sudden ultimatum, etc.) which such a stockpile gives to the state in question, and not in terms of money. In this situation, the rate at which these weapons are amassed may be very important to the evader. He may have to make political concessions that are distasteful to him until he has a sufficient stockpile. Also, in time, an effective antimissile missile may be developed by his opponent, making his missiles obsolete. Exponential discounting may be useful as a device for approximating the evader's interest in quick returns.

One must bear in mind that when any realistic situation is modeled, many assumptions and approximations are usually necessary before the problem can be simplified to the point where it can be handled analytically. When rewards do not involve money, the problem of establishing a set of utilities for the various possible events usually poses far more difficulties than does the problem of finding an appropriate discount factor. In the above example, the assumption that the utility of a stockpile of missiles is proportional to the number of missiles (an automatic result of the model) is far more open to criticism than is the discounting device.

In this chapter we shall use the reward coefficients ρ_i, η_i, τ_i and λ_i that were defined in Sec. 6.1. In addition, let us define

α = interest rate,

$d = e^{-\alpha}$ = discount factor per unit time,

$\gamma_i = \frac{1-d^{\tau_i}}{\alpha}$ = effective search time for box i.

When the evader hides in box i and the searcher looks into box j ($j \neq i$), the evader receives income at a rate ρ_i for τ_j units of time. The reward of this event, which is equal to the utility that applies when the event begins, is

$$\int_0^{\tau_j} \rho_i e^{-\alpha t} dt = \rho_i \left(\frac{1-e^{-\alpha \tau_j}}{\alpha} \right)$$

$$= \rho_i \gamma_j \quad .$$

If the evader is hiding in box j, the reward is $(\rho_j - \eta_j)\,\gamma_j$. This is equivalent to the $\alpha = 0$ case where the search time for box j is γ_j. As a result, γ_j is called the effective search time. Both τ_j, the actual search time, and γ_j will be used in our equations.

A final detection loss (or negative reward) can be incurred by the evader when he is found. This loss may be used to account for a penalty that the evader incurs as a result of being found and also for the loss of earnings, in the expected sense, that can result if he can be found sometime during the look rather than just at the end. The utility of this loss when it is evaluated at the end of the final look will be represented by λ_i.

Now that time is an important consideration, we must also consider the time during which moving can occur. Usually, we can expect that a dead time between looks provides the opportunity for this action. It is convenient to let the search time τ_i include the dead time at the end of the look. This may require some adjustments in the various earning rates, and so forth, but usually there is no reason for assuming that the evader cannot continue earning during the dead time. In fact, the moving cost μ_i may result partially from the loss in earnings which occurs during the time required for the move. The value of each μ_i will be defined as the loss in utility that applies just before the next look and hence at the end of the time available for moving. This definition allows us to write our equations as though moving occurred instantaneously.

The introduction of discounting into the search evasion game does not appreciably affect its general behavior but does increase the notational complexity of some of the equations. As a result, we shall merely paraphrase the developments of Chapters 2 through 5 as we did in Chapter 6 when the generalized reward structure was introduced. Any equation that must be modified will again bear its original number, but this time the double section sign will be attached. Also, any changes in the game's properties that affect the previous results will be discussed.

7.2 INADMISSIBLE BOXES

Perhaps the chief phenomenon that is introduced by discounting and that must be considered before continuing concerns inadmissible boxes. In the previous chapters, we found that the evader should always hide in either box with a nonzero probability and that the searcher's good strategy always requires at least one look into each of them. This is true even if one box has a much lower detection probability q or a much higher earning rate ρ. The evader's good strategy requires P to be unequal to zero or one, since otherwise the searcher, if he knew this strategy, could always look into the correct box. Similarly, the searcher's good strategy always results in some looks into each box, for otherwise the evader could receive an infinite payoff. When discounting applies, however, the evader can never receive an infinite payoff. As a result, we may find that the detection probabilities, earning rates, and so forth, are biased so much in the favor of one box that the inferior one is not used by either player. If this occurs, the box is inadmissible.

The conditions under which a box is inadmissible can easily be found. To do this, first consider the case where the evader hides and remains in box i and the searcher always looks into box j. In this situation, the evader has an earning rate ρ_i that continues for all time. Therefore, he receives a total payoff equal to ρ_i/α. If, on the other hand, he were to hide in box j until he were found, while the searcher always looked there, he would receive a payoff equal to

$$U = (\rho_j - \eta_j)\, \gamma_j - q_j d^{\tau_j} \lambda_j + r_j d^{\tau_j} U \quad ,$$

or

$$U = \frac{(\rho_j - \eta_j)\, \gamma_j - q_j d^{\tau_j} \lambda_j}{1 - r_j d^{\tau_j}} \quad .$$

Thus, if

$$\frac{\rho_i}{\alpha} \leqslant \frac{(\rho_j - \eta_j)\, \gamma_j - q_j d^{\tau_j} \lambda_j}{1 - r_j d^{\tau_j}} \quad ,$$

box i is inadmissible, for the evader would be foolish to hide in box i even if the searcher always looked into box j. Similar reasoning shows that the searcher should never look into an inadmissible box unless he knows that the evader is foolishly hiding there with a sufficient nonzero probability. Since we require each ρ, η, τ, and λ to be nonnegative, both boxes cannot be inadmissible.

If one of the boxes is inadmissible, the two-box game loses all interest, for the game degenerates to a trivial one-box game. It should be clear that the inadmissibility condition does not depend upon the moving costs or, in fact, on whether moving is allowed or not. In the remaining sections of this chapter, we shall assume that neither box is inadmissible.

7.3 G^{∞}: THE NO-MOVE GAME

When moving is not allowed, the modified game F^{∞} may again be used. The payoff function $U^{\infty}(P)$ has the same properties of being continuous, and convex. Perhaps the only difference in its general appearance worthy of note is that the magnitude of the slope of this function no longer becomes arbitrarily large as P approaches zero and one. This follows from the fact that the payoff is no longer infinite if the evader hides in one box and the searcher always looks into the other. The fundamental recursion equation that now applies is

$$U^{\infty}(P) = \min \begin{cases} U^{\infty}(P;1) = P\left[(\rho_1 - \eta_1)\,\gamma_1 - q_1 d^{\tau_1}\lambda_1\right] + (1-P)\,\rho_2\gamma_1 \\ \qquad + d^{\tau_1}[Pr_1 + 1 - P]\,U^{\infty}\left[\dfrac{Pr_1}{Pr_1 + 1 - P}\right] \\ U^{\infty}(P;2) = P\rho_1\gamma_2 + (1-P)\left[(\rho_2 - \eta_2)\,\gamma_2 - q_2\, d^{\tau_2}\lambda_2\right] \\ \qquad + d^{\tau_2}[P + (1-P)\,r_2]\,U^{\infty}\left[\dfrac{P}{P + (1-P)\,r_2}\right] \end{cases} . \qquad (2\text{-}4)\S\S$$

There again exists a P_0 (see Appendix A) where the searcher should look into box 1 if P is greater than P_0, and into box 2 if it is less than P_0. The value of P_0 can be found in the usual way and is

$$P_0 = \frac{1/\beta_1}{1/\beta_1 + 1/\beta_2} \quad ,$$

where

$$\beta_i = \frac{d^{\tau_i} q_i}{\gamma_i}\,(\rho_i + \alpha\lambda_i) + \alpha\eta_i \quad .$$

Perhaps a simpler way of expressing this rule now that the above expression is so complex is simply to state that the searcher should look into that box for which the associated expression

$$p_i \left[\frac{d^{\tau_i} q_i}{\gamma_i}\,(\rho_i + \alpha\lambda_i^{d}) + \alpha\eta_i\right] = p_i\beta_i$$

is the larger. Note that if α is very small, $\gamma_i \cong \tau_i$, $d^{\tau_i} \cong 1$, and P_0 is approximately equal to $(\tau_1/\rho_1 q_1)/(\tau_1/\rho_1 q_1 + \tau_2/\rho_2 q_2)$ as in Chapter 6.

Every term in the expression for P_0 is positive. Therefore, $0 < P_0 < 1$. This occurs even if one of the boxes is inadmissible. This should not be surprising since, even when a box is inadmissible, the searcher should look there if he knows that the probability that the evader is there is sufficiently high. On the other hand, $U^{\infty}(P)$ will be a maximum at $P = 0$ if box 1 is inadmissible and at $P = 1$ if box 2 is.

The recurrent region (P_{01}, P_{02}) has the same properties as in Chapter 6. When $r_1^{n_1} = r_2^{n_2}$, the search sequence in this region is again periodic. If $\log r_2/\log r_1$ is irrational, it can be approximated by n_1/n_2 as before. With discounting, this approximation can be looser than before, since the effects of incorrect looks in the distant future are ameliorated by the discount factor as well as by the decreasing probability of survival. With a given pair of integers n_1 and n_2, a chain diagram can be devised in the same manner as before. Transient states can be added when necessary, as was discussed in Chapter 6.

The functional relationship between the payoff associated with a given state in the searcher's chain diagram and the one into which it is transformed by the next look is

$$s_i \xrightarrow{1} s_j \Longrightarrow$$

$$a_i = (\rho_1 - \eta_1)\,\gamma_1 - q_1 d^{\tau_1}\lambda_1 + r_1 d^{\tau_1} a_j \quad ,$$

$$b_i = \rho_2\gamma_1 + d^{\tau_1} b_j \quad ;$$

$$s_i \xrightarrow{2} s_j \Longrightarrow$$

$$a_i = \rho_1\gamma_2 + d^{\tau_2} a_j \quad ,$$

$$b_i = (\rho_2 - \eta_2)\,\gamma_2 - q_2 d^{\tau_2}\lambda_2 + r_2 d^{\tau_2} b_j \quad . \tag{2-11}§§$$

If one wishes to express the payoff associated with a state s_i in terms of that associated with a state s_j when a sequence of looks transforms s_i into s_j, our notation must be redefined slightly. In particular, the sequence must be defined by a set $\{\tau_m(n)\}$ where $\tau_m(n)$ is the time at which the n^{th} look into box m is completed. We can again let k_m represent the total number of looks into box m. Furthermore, it is convenient to let τ_t represent the total time of the sequence. Clearly, $\tau_t = \max\{\tau_1(k_1), \tau_2(k_2)\}$, and we find that

$$s_i \xrightarrow{\{\tau_m(n)\}} s_j \Longrightarrow$$

$$a_i = q_1 \sum_{n=1}^{k_1} r_1^{n-1} \left[\rho_1 \left(\frac{1 - d^{\tau_1(n)}}{\alpha}\right) - \eta_1\gamma_1 d^{-\tau_1} \sum_{k=1}^{n} d^{\tau_1(k)} - \lambda_1 d^{\tau_1(n)}\right]$$

$$+ r_1^{k_1} \left[\rho_1 \left(\frac{1 - d^{\tau_t}}{\alpha}\right) - \eta_1 d^{-\tau_1} \sum_{n=1}^{k_1} d^{\tau_1(n)} + d^{\tau_t} a_j\right] \quad ,$$

$$b_i = q_2 \sum_{n=1}^{k_2} r_2^{n-1} \left[\rho_2 \left(\frac{1 - d^{\tau_2(n)}}{\alpha} \right) - \eta_2 \gamma_2 d^{-\tau_2} \sum_{k=1}^{n} d^{\tau_2(k)} - \lambda_2 d^{\tau_2(n)} \right]$$

$$+ r_2^{k_2} \left[\rho_2 \left(\frac{1 - d^{\tau_t}}{\alpha} \right) - \eta_2 d^{-\tau_2} \sum_{n=1}^{k_2} d^{\tau_2(n)} + d^{\tau_t} b_j \right] \qquad \text{(2-13)§§}$$

These equations are rather complex and it may prove simpler to compound Eq. (2-11)§§ if the sequence is fairly short. In Eq. (2-13)§§, the payoff associated with each of the possible times at which detection can occur is found by first calculating the utility contributed by the earning rate ρ, then subtracting the loss in earnings from each of the looks into the correct box, and finally deducting the detection loss. These equations could, of course, be formulated in many other ways. Perhaps the main reason for doing it this way is that it carries over fairly directly to the many-box case. When s_i is transformed into itself by a sequence of looks, the coefficients a_i and b_i can be expressed in closed form by the usual extension of (2-13)§§.

Once the value of P at which $U^{\infty}(P)$ is a maximum and the payoffs associated with the search states optimum at this point have been found, the searcher's good strategy can be completed in the usual manner.

7.4 GAME G°: $\mu_1, \mu_2 = 0$

This game may be solved in exactly the same manner as it was in Chapter 6 once the effects of discounting have been introduced into the necessary equations. Now, however, we must require that neither box be inadmissible. The evader's good strategy can be obtained from the solution of the equations

$$U°(P_0) = U°(P_0; 1) = \frac{P_0\left[(\rho_1 - \eta_1)\,\gamma_1 - q_1 d^{\tau_1} \lambda_1\right] + (1 - P_0)\,\rho_2 \gamma_1}{1 - d^{\tau_1}(1 - P_0 q_1)}$$

$$= U°(P_0; 2) = \frac{P_0 \rho_1 \gamma_2 + (1 - P_0)\left[(\rho_2 - \eta_2)\,\gamma_2 - q_2 d^{\tau_2} \lambda_2\right]}{1 - d^{\tau_2}[1 - (1 - P_0)\,q_2]} ,$$

where $0 \leqslant P_0 \leqslant 1$. The searcher's good strategy can be obtained from

$$W(Y_0) = W(Y_0; 1) = \frac{Y_0\left[(\rho_1 - \eta_1)\,\gamma_1 - q_1 d^{\tau_1} \lambda_1\right] + (1 - Y_0)\,\rho_1 \gamma_2}{1 - Y_0 r_1 d^{\tau_1} - (1 - Y_0)\,d^{\tau_2}}$$

$$= W(Y_0; 2) = \frac{Y_0 \rho_2 \gamma_1 + (1 - Y_0)\left[(\rho_2 - \eta_2)\,\gamma_2 - q_2 d^{\tau_2} \lambda_2\right]}{1 - Y_0 d^{\tau_1} - (1 - Y_0)\,r_2 d^{\tau_2}} ,$$

where $0 \leqslant Y_0 \leqslant 1$. Appendix D shows that these solutions exist and that $U°(P_0) = W°(Y_0) = V°$, the value.

The above equations will not, of course, yield a solution if one of the boxes is inadmissible. When this occurs, the two curves $U°(P;1)$ and $U°(P;2)$ still intersect at some point P_0 in the interior of the interval (0, 1). Hence, there still exists a strategy for the evader that yields a payoff independent of the searcher's strategy. This is not the evader's good strategy, however, for he can guarantee a larger payoff by hiding in the admissible box with probability one; that is, $U°(P_0;1) = U°(P_0;2) < \max_P \{\min_i U°(P;i)\}$. It should not be surprising to find, in contrast, that the two curves $W°(Y_0;1)$ and $W°(Y_0;2)$ do not intersect in (0, 1) when a box is inadmissible. If they did, the associated strategy would produce a payoff independent of the evader's strategy. In general, when each player has a strategy that yields a payoff entirely independent of the other's, the payoffs must be equal and the strategies must be good strategies.

7.5 GAME G

When one or both of the moving costs are no longer equal to zero, the techniques developed in Chapters 4 and 5 may again be used once the appropriate changes have been introduced into the various equations. The fundamental functional equations for the modified games F and F' are now

$$U'(P) = \min \begin{cases} U'(P;1) = P\left[(\rho_1 - \eta_1)\,\gamma_1 - q_1 d^{\tau_1}\lambda_1\right] + (1-P)\,\rho_2\gamma_1 \\ \qquad + d^{\tau_1}[Pr_1 + 1 - P]\, U\left[\dfrac{Pr_1}{Pr_1 + 1 - P}\right] \\ U'(P;2) = P\rho_1\gamma_2 + (1-P)\left[(\rho_2 - \eta_2)\,\gamma_2 - q_2 d^{\tau_2}\lambda_2\right] \\ \qquad + d^{\tau_2}[P + (1-P)\,r_2]\, U\left[\dfrac{P}{P + (1-P)\,r_2}\right] \end{cases} \qquad (4\text{-}3)\S\S$$

and

$$U(P) = \max_{P'} \begin{cases} -\mu_1(P' - P) + U'(P') \quad , \quad P' \geqslant P \\ -\mu_2(P - P') + U'(P') \quad , \quad P' \leqslant P \quad . \end{cases} \qquad (4\text{-}4)\S$$

The functions $U(P)$ and $U'(P)$ again have the same basic properties that allow the previous solution techniques to be used. Both functions are continuous and convex. In general, they will be piecewise linear if the moving costs are not prohibitive. In game F' there exists a P_0 where $0 < P_0 < 1$ that has the usual properties. The proof that these properties are still satisfied is found in Appendix B.

The moving region (P_-, P_+) is again defined by Eq. (4-5)§. The moving costs are prohibitive if they both satisfy Eq. (4-7)§. When this occurs, both $U(P)$ and $U'(P)$ are again identical to $U^\infty(P)$ over the no-move region and also over the recurrent region, which it contains under these conditions. As we have seen, once the moving costs become prohibitive, the searcher's good strategy becomes identical to that in G^∞. The evader should never move as long as the searcher uses this good strategy. We also found that the no-move regions never completely disappeared as long as μ_1 and μ_2 were finite (unless q_1 or $q_2 = 1$). It was necessary to calculate the values of P_- and P_+

if we wished to obtain the evader's complete good strategy. These bounding points of the no-move region are of use to the evader when the searcher uses an inadmissible sequence that transforms P into a moving region.

With one exception, all of these properties still hold when discounting is considered. The one exception is that the moving regions can now disappear completely when μ_1 and μ_2 are finite. This can occur because the magnitude of the slope of $U^{\infty}(P)$ no longer approaches infinity as P approaches zero or one. In game F^{∞}, the searcher should always look into box 1 if P = 1. Therefore,

$$\lim_{P\to 1} U^{\infty}(P) = \left[\frac{(\rho_1 - \eta_1)\,\gamma_1 - q_1 d^{\tau_1}\lambda_1}{1 - r_1 d^{\tau_1}}\right] P + \frac{\rho_2}{\alpha}(1 - P) \quad .$$

Similarly,

$$\lim_{P\to 0} U^{\infty}(P) = \frac{\rho_1}{\alpha}\,P + \left[\frac{(\rho_2 - \eta_2)\,\gamma_2 - q_2 d^{\tau_2}\lambda_2}{1 - r_2 d^{\tau_2}}\right](1 - P) \quad .$$

It follows that if

$$\mu_1 > \frac{\rho_1}{\alpha} - \frac{(\rho_2 - \eta_2)\,\gamma_2 - q_2 d^{\tau_2}\lambda_2}{1 - r_2 d^{\tau_2}} \quad ,$$

and

$$\mu_2 > \frac{\rho_2}{\alpha} - \frac{(\rho_1 - \eta_1)\,\gamma_1 - q_1 d^{\tau_1}\lambda_1}{1 - r_1 d^{\tau_1}} \quad ,$$

U(P) and U'(P) will be identical to $U^{\infty}(P)$ over the entire interval (0, 1). Under these conditions, the evader should never move. Of course, as μ_1 and μ_2 increase from zero, they will become prohibitive before both of the above conditions occur. The above bounds, however, are easily calculated and may possibly indicate that the moving costs are definitely prohibitive when μ_1 and μ_2 are very large. Furthermore, they show that when α is unequal to zero the moving cost will be prohibitive for sufficiently large but finite moving costs even when one, but not both, of the detection probabilities is equal to one.

When the moving costs are not prohibitive, games F and F' may be solved by going through the same process of studying the manner in which the optimum chain diagram changes from strategy interval to strategy interval as μ_1 and μ_2 increase in constant ratio. Also, the usual techniques may be used to calculate U(P) and U'(P) once the correct chain diagram has been found. To avoid repetition, only those equations in Chapter 6 that must be changed will be listed here.

Equation (4-8)§ could be used in that chapter to express the payoff of a state s_i in terms of the payoff of some other state s_j when the transformation of s_i into s_j did not involve any move transitions. Thus, it could be used to express $U'_+(P)$ in terms of $U_-(P)$, and so forth. This equation is identical to Eq. (2-13)§, which was used in the no-move game. It must now be replaced by Eq. (2-13)§§.

Equation (4-12)§ must also be rewritten and is now

$$P_0\left[(\rho_1-\eta_1)\,\gamma_1-q_1d^{\tau_1}\lambda_1+r_1d^{\tau_1}a_-\right]+(1-P_0)\left[\rho_2\gamma_1+d^{\tau_1}b_-\right]=P_0\left[\rho_1\gamma_2+d^{\tau_2}a_+\right]$$

$$+(1-P_0)\left[(\rho_2-\eta_2)\,\gamma_2-q_2d^{\tau_2}\lambda_2+r_2d^{\tau_2}b_+\right] \quad . \tag{4-10)§§}$$

This equation is used in the above form when P_{01} belongs to the interior of π_- and P_{02} to the interior of π_+. As in Chapter 4, the coefficients a_- and b_- or a_+ and b_+ must be replaced by those appropriate when only one of the moving regions extends into the recurrent region (see Sec.4.7.1). Equations (4-9)§, (4-10) and (4-11) can be used without any alterations since discounting has no effect on them.

Once games F and F' have been solved and the evader's good strategy in G has been found, the searcher's good strategy can also be obtained in the usual manner. The functional equations of games H and H' are now

$$\begin{aligned}W_i'(P) = y_i(1) &\left\{\begin{aligned}&P\left[(\rho_1-\eta_1)\,\gamma_1-q_1d^{\tau_1}\lambda_1\right]+(1-P)\,\rho_2\gamma_1\\ &+d^{\tau_1}[Pr_1+1-P]\,W_{i|1}\left[\frac{Pr_1}{Pr_1+1-P}\right]\end{aligned}\right\}\\ +\,y_i(2)&\left\{\begin{aligned}&P\rho_1\gamma_2+(1-P)\left[(\rho_2-\eta_2)\,\gamma_2-q_2d^{\tau_2}\lambda_2\right]\\ &+d^{\tau_2}[P+(1-P)\,r_2]\,W_{i|2}\left[\frac{P}{P+(1-P)\,r_2}\right]\end{aligned}\right\}\end{aligned} \tag{5-1)§§}$$

and

$$W_i(P)=\begin{cases}-\mu_2P+W_i'(0) & , \quad \dfrac{dW_i'(P)}{dP}<-\mu_2\\ W_i'(P) & , \quad -\mu_2\leqslant\dfrac{dW_i'(P)}{dP}\leqslant\mu_1\\ -\mu_1(1-P)+W_i'(1) & , \quad \dfrac{dW_i'(P)}{dP}>\mu \quad .\end{cases} \tag{5-2)§}$$

The correct transition diagram can be derived from the searcher's chain diagram of game F', once that game has been solved, in the usual manner. The only computational changes required in calculating the probability distributions of the mixed states are those which result from the new form of Eq. (5-1)§§. The reasoning used in Chapter 5 and Appendix C to show that the searcher could indeed limit the evader to U(P) when the initial P is known is still valid in the discounting case. To complete the searcher's good strategy, the starting rule $\underline{Y}_0$ can be computed precisely as in Chapter 5. Discounting has no effect on this computation. No look, hence no time, is involved in the selection of a starting state for the Markov process that generates the search sequence.

This completes the discussion of the two-box search evasion game. As we have seen, all of the important properties of this game occur with the simple reward structure used in Chapters 2 through 5. Discounting, of course, raises the interesting possibility that a box may be inadmissible. This occurs only in extremely biased cases, however, when the interest rate is fairly high. The various equations become more complex algebraically when the generalized reward structure and discounting are introduced. On the other hand, we have seen that the same general computational methods are still valid, that the number of equations necessary for a given calculation does not increase, and that no new nonlinearities (except in G^*) arise. Thus, the increase in computational complexity is not great and is a small price to pay for the increase in generality achieved in these last two chapters.

CHAPTER 8
THE SEARCH EVASION GAME WITH N BOXES

8.1 INTRODUCTION

In the previous chapters, the two-box search evasion game has been considered in some detail, and it is now appropriate to turn our attention to the more general N-box game. As would be expected, the behavior of the game becomes more complex when three or more boxes are involved. We shall first examine the limiting games G^{∞} and G°. These games behave much as before and only the computational effort becomes more involved. The good search strategy associated with G° will be of particular interest since it may always be used to limit the evader to V°, the value of G°, when the moving costs are unequal to zero or when evasive countermeasures other than moving are available to the evader.

When game G is considered, we shall find that some of the properties fundamental to the solution techniques of the previous chapters no longer hold. For example, the searcher's good strategy can no longer be generated by a simple Markov process, and there no longer exists a finite number of strategy intervals as the moving costs increase in constant ratio from zero up to a point where they are all prohibitive. As a result, no general method for solving G when there are more than two boxes has been developed. A simple example has been solved, however, and will be used to illustrate some of the problems that can be expected in the search for exact solution techniques. It will also indicate the extreme magnitude of the computational effort that could be expected if a general method were devised and, therefore, the desirability of finding an efficient method for obtaining strategies that limit the evader to a payoff close to the value. A particular approach will be suggested for future research.

The reward structure that will be used for the N-box game is the same as that used in Chapter 6. Thus,

ρ_i = the evader's earning rate when he hides in box i and the searcher looks elsewhere,

η_i = the loss in earning rate when the searcher looks into the correct box,

τ_i = the time required to examine box i,

λ_i = the detection loss.

We again require that ρ_i, $\tau_i > 0$ and η_i, $\lambda_i \geqslant 0$. If discounting is used, we can again let

α = compound interest rate,

$d = e^{-\alpha}$ = discount factor per unit time,

$$\gamma_i = \frac{1 - d^{\tau_i}}{\alpha} = \text{effective search time.}$$

8.2 GAME G^{∞}

In G^{∞}, the evader cannot move between looks and the N-box case is quite similar to the two-box version. A strategy for the evader consists of the selection of a probability vector $\underline{P} = \{p_1, p_2, \ldots, p_N\}$ that is defined over a bounded $N - 1$ space (since $\sum_{i=1}^{N} p_i = 0$, $p_i \geqslant 0$). A pure strategy for the searcher consists of an infinite search sequence that is used as long as necessary. The probability space over which $\underline{P}$ is defined can best be represented by a regular simplex of degree $N-1$ and barycentric coordinates. In the three-box case the simplex is an equilateral triangle and in the four-box case a regular tetrahedron. For each coordinate p_i, there is an associated vertex or extreme point of the simplex where $p_i = 1$ and an opposite face over which $p_i = 0$. This face is the regular simplex of one lower degree that is generated by all of the remaining vertices. At any given point within the simplex, the value of p_i is equal to the distance from this point to the i^{th} face. Requiring the altitude of the simplex to equal one insures that $\sum_{i=1}^{N} p_i = 1$ for any $\underline{P}$ belonging to it.

In the modified game F^{∞}, the searcher is informed of the initial position of $\underline{P}$ and can calculate its a posteriori position after each unsuccessful look. Thus, if the searcher looks into box i, we find that

$$\underline{P} \xrightarrow{i} \underline{P}' \Longrightarrow$$

$$p_i' = \frac{p_i r_i}{1 - p_i q_i} \quad ,$$

$$p_j' = \frac{p_j}{1 - p_i q_i} \quad , \quad i \neq j \quad . \tag{8-1}$$

A sequence of looks that involves a total of k_i looks into box i for each i transforms $\underline{P}$ according to

$$\underline{P} \xrightarrow{\{k_i\}} \underline{P}' \Longrightarrow$$

$$p_i' = \frac{p_i r_i^{k_i}}{\sum_{j=1}^{N} p_j r_j^{k_j}} \quad . \tag{8-2}$$

As before, the order of the sequence has no effect on the final transformation.

Given a particular $\underline{P}$, the searcher must decide where to look next. Since $\underline{P}$ is transformed by the search process only, an optimum infinite search sequence can be associated with each $\underline{P}$. As before, the payoff associated with any arbitrary sequence is linear in $\underline{P}$. The payoff function $U^{\infty}(\underline{P})$, which results when an optimum sequence is used for the $\underline{P}$ in question, is formed by the lower bound on the ensemble of payoffs generated by all infinite search sequences. The function $U^{\infty}(P)$ must be continuous and convex. Furthermore, it may be piecewise linear in the interior of the simplex over which $\underline{P}$ is defined. That is, the simplex may be partitioned into a set of hypervolumes within each of which $U^{\infty}(\underline{P})$ is linear in $\underline{P}$. Over each of these hypervolumes, a particular infinite search sequence is optimum. As $\underline{P}$ approaches any boundary, these hypervolumes must become arbitrarily small.

The functional equation that defines the optimum payoff function is now

$$U^{\infty}(\underline{P}) = \min_i \{U^{\infty}(\underline{P};i)\} \quad ,$$

where

$$U^{\infty}(\underline{P}; i) = \gamma_i \sum_{k=1}^{N} p_k \rho_k - p_i\left(\gamma_i \eta_i + q_i d^{\tau_i} \lambda_i\right) + d^{\tau_i}(1 - p_i q_i)\, U^{\infty}(\underline{P}') \quad , \tag{8-3}$$

and

$$\underline{P} \xrightarrow{i} \underline{P}' \quad .$$

The optimum search strategy for F^∞ can be derived heuristically in much the same manner as before. It is again convenient to let $U^\infty(\underline{P}; ij)$ represent the payoff that results if a look into box i is followed by a look into box j and then by an optimum sequence. It follows from Eq. (8-3) that

$$\begin{aligned} U^\infty(\underline{P}; ij) = \gamma_i \sum_{k=1}^{N} \rho_k p_k - p_i \left(\gamma_i \eta_i + q_i d^{\tau_i} \lambda_i\right) \\ + d^{\tau_i} \left[\gamma_j \sum_{k=1}^{N} \rho_k p_k - p_i q_i \gamma_j \rho_i - p_j \left(\gamma_j \eta_j + q_j d^{\tau_j} \lambda_j\right)\right] \\ + d^{\tau_i + \tau_j} (1 - p_i q_i - p_j q_j)\, U^\infty(\underline{P}') \quad , \end{aligned}$$

where

$$\underline{P} \xrightarrow{ij} \underline{P}' \quad .$$

If we set $U^\infty(\underline{P}; ij)$ equal to $U^\infty(\underline{P}; ji)$, the term $U^\infty(\underline{P}')$ cancels, and we find that

$$p_i \beta_i = p_j \beta_j \quad ,$$

where

$$\beta_i = \frac{d^{\tau_i} q_i}{\gamma_i} (\rho_i + \alpha \lambda_i) + \alpha \eta_i \quad .$$

When $\alpha = 0$,

$$\beta_i = \frac{\rho_i}{\tau_i} q_i \quad .$$

The equation $p_i\beta_i = p_j\beta_j$ defines a hyperplane of degree $N - 2$ that intersects the line joining vertices i and j and also all of the remaining vertices. It therefore partitions the simplex into two parts. In the space where $p_i\beta_i > p_j\beta_j$, the sequence ij + optimum is preferable to ji + optimum, and so forth.

This does not imply that either of these sequences is necessarily the optimum one. On the other hand, it is not unreasonable to assume that the optimum strategy will require a look into box i before a look into box j when this occurs. Carrying this reasoning a little further, we should expect the optimum search rule to require the next look to be into that box for which $p_i\beta_i$ is a maximum.

The above argument does not, of course, prove that this is indeed the optimum search rule. It has provided a convenient means for deriving the form of the expression β_i, however, and with this expression it is not too difficult to prove that the above search rule is indeed the optimum one. The proof is contained in Appendix A. Since the optimum search rules of F^∞ developed previously for the two-box game are special cases of this rule, this proof also establishes their validity.

As a result of the simplicity of the optimum search rule – the searcher should merely look into a box for which $p_i\beta_i$ is a maximum – many interesting properties can be developed. In order to do this, we must first examine more closely the behavior of the state vector $\underline{P}$ as a function of a sequence of unsuccessful looks. For any given $\underline{P} = \{p_i\}$, we may define an associated set $\{\xi_i\}$ which satisfies the equations $p_1\xi_1 = p_2\xi_2 = \ldots = p_N\xi_N$ such that $\xi_i \geqslant 0$ for all i. Such a set is not unique unless it is normalized, but any set of this form will uniquely determine a $\underline{P}$ belonging to the probability simplex. An equation of the form $p_i\xi_i = p_j\xi_j$ defines a hyperplane that intersects all but the i^{th} and j^{th} vertices and, in addition, the line joining these remaining two. Since $\underline{P}$ must belong to all of these hyperplanes, it lies at their common point of intersection. Although there are $N(N-1)$ such hyperplanes, only $N-1$ are independent and any arbitrary set $\{\xi_i\}$ will have a unique point of intersection.

The most interesting property of these hyperplanes concerns the way in which $\underline{P}$ is transformed from one to another by a sequence of unsuccessful looks. If $\underline{P}$ is defined by the set $\{\xi_i\}$, the a posteriori $\underline{P}'$ resulting from an unsuccessful look into box k must satisfy the equations

$$\frac{p'_k\xi_k}{r_k} = p_i\xi_i \quad , \quad i \neq k \quad ;$$

$$p_j\xi_j = p_i\xi_i \quad , \quad i, j \neq k \quad .$$

If $\underline{P}$ originally belongs to the hyperplane $p_i\xi_i = p_j\xi_j$, it will not leave it until either box i or box j is examined. If an arbitrary sequence includes k_i looks into box i and k_j looks into box j, the a posteriori $\underline{P}'$ will belong to the hyperplane

$$\frac{p'_i\xi_i}{r_i^{k_i}} = \frac{p'_j\xi_j}{r_j^{k_j}} \quad .$$

This condition applies even if the arbitrary sequence includes looks into other boxes.

For a given pair (i, j), all hyperplanes of the form $p_i\xi_i = p_j\xi_j$, or (ξ_i, ξ_j) for short, intersect the line joining the i^{th} and j^{th} vertex, and they can be ordered by their intersection along this line. If $(\xi'_i/\xi'_j) < (\xi_i/\xi_j)$, the hyperplane (ξ'_i, ξ'_j) intersects this line at a point closer to the i^{th} vertex (where $p_i = 1$), and we can say that $(\xi'_i\xi'_j)$ lies on the i^{th} side of (ξ_i, ξ_j). Similarly, the vector $\underline{P}$ lies to the i^{th} side of (ξ_i, ξ_j) if $(p_j/p_i) < (\xi_i/\xi_j)$.

Any vector that lies on the i^{th} side of (β_i, β_j) will remain there until box i is examined, and the searcher's optimum strategy will require at least one look into box i before box j is examined for the first time. Therefore, by ordering of the terms $p_i\beta_i$ in decreasing magnitude for a given $\underline{P}$, we can tell more about the associated optimum sequence than merely which box should be examined first.

As an illustration of the manner in which the simplex is partitioned into a set of hypervolumes over each of which a particular next look is optimum, let us consider the three-box case. Here, the simplex is an equilateral triangle, and a hyperplane (ξ_i, ξ_j) is a line joining the k^{th} $(k \neq i, j)$ vertex to the line connecting the i^{th} and the j^{th}. A hyperplane (β_i, β_j) partitions this triangle into two parts. If $\underline{P}$ lies to one side, box i will be examined at least once before box j is, and so forth. The simplex is of the form shown in Fig. 23. The part of each hyperplane (β_i, β_j) over which the first look can be made into either box i or j is indicated by a solid line, and the three solid-line segments of this type partition the triangle into the three areas in which a particular next look is optimum. If $\underline{P}$ belongs to the broken section of the line (β_i, β_j), the next optimum look will be made into some other box, i.e., the remaining one. As long as $0 < p_i, p_j < 1$, both probabilities will be increased in constant ratio by such a look. Eventually, such a point will be transformed into the solid section by an optimum sequence of looks into the other box and a look into box i or box j will then be optimum.

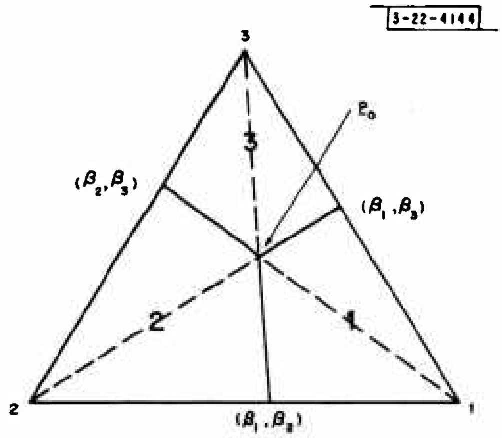


Fig. 23. The three-box simplex: the optimum next look.

At the point $\underline{P}_0$, where the three hyperplanes intersect, a look into any box is optimum, and each box should be examined once during the first three looks. In the more general N-box case, any of the N! possible orderings of one look into each box is optimum during the first N looks.

A recurrent region can be defined for the N-box game. This region consists of the minimum hypervolume from which no $\underline{P}$ belonging to it can be removed by an optimum search sequence (of <u>unsuccessful</u> looks) and into which any other $\underline{P}$ not belonging to a boundary of the simplex must eventually be transformed. We shall first consider the form of this region when all the detection probabilities are less than one. When $\underline{P}$ belongs to the hyperplane (ξ_i, ξ_j) it can be transformed only to the j^{th} side of it by a look into box i. For a look into box i to be optimum, however, $\underline{P}$ must belong to or lie to the i^{th} side of the hyperplane (β_i, β_j). It follows that no $\underline{P}$ can be transformed to the j^{th} side of the hyperplane $[(\beta_i/r_i), \beta_j]$ by an optimum look. Similarly, we see that no $\underline{P}$ can be transformed to the i^{th} side of the hyperplane $[(\beta_i, (\beta_j/r_j)]$ by an optimum look once it lies on or to the j^{th} side of it. Therefore, the hypervolume

$$\frac{\beta_j r_i}{\beta_i} \leqslant \frac{p_i}{p_j} \leqslant \frac{\beta_j}{\beta_i r_j}$$

must contain the recurrent region. In fact, the recurrent region must clearly consist of that hypervolume which satisfies this requirement for all pairs (i, j). The form of the recurrent region,for the three-box game (Fig. 24) is an irregular hexagon.

Since the recurrent region is convex and bounded by a set of linear hyperplanes, it can be generated by a set of extreme points. It can be shown that these extreme points are the $2^N - 2$ possible points into which $\underline{P}_0$ can be transformed during the first $N - 1$ looks of an optimum sequence.

If there exists a set $\{n_i\}$ such that $r_1^{n_1} = r_2^{n_2} = \ldots = r_N^{n_N}$, the optimum search sequence will be periodic for any P belonging to the recurrent region. A set can always be found that satisfies

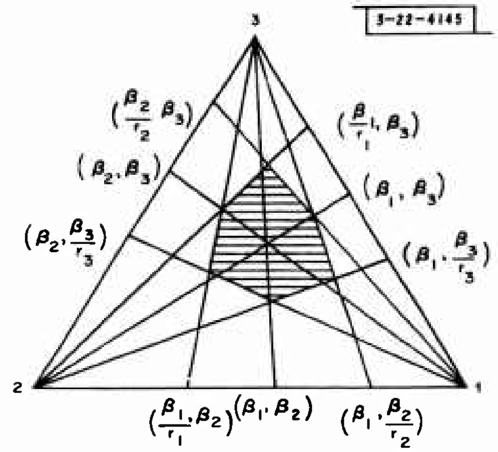


Fig. 24. The recurrent region in the three-box game.

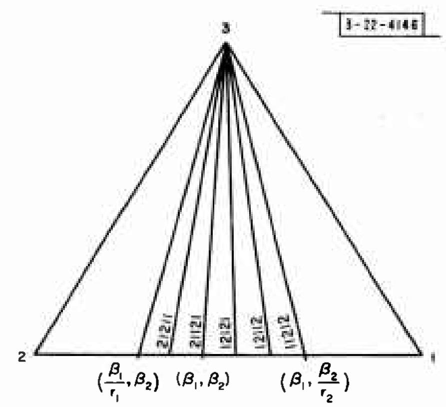


Fig. 25. Relative ordering of looks into boxes 1 and 2 ($r_1^3 = r_2^2$).

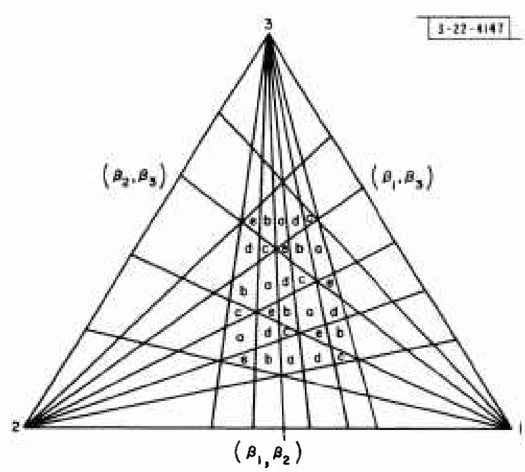


Fig. 26. General form of the recurrent region ($r_1^3 = r_2^2 = r_3$).

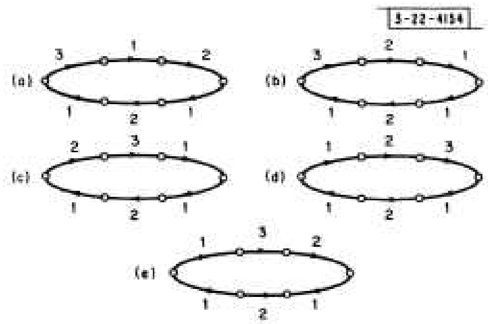
the above equations to any desired degree of accuracy as long as each q_i is less than one. In each period of such a sequence, each box i will be examined n_i times. A set of hyperplanes partitions this region into a set of hypervolumes, each having a different sequence of this type associated with it. The form of this partition can be found once we note that the relative ordering of looks into boxes i and j within the optimum sequence is invariant over a hyperplane of the form (ξ_i, ξ_j). A set of hyperplanes of this form must, therefore, partition the recurrent region into a set of hypervolumes, over each of which the relative order of looks into boxes i and j is unique and periodic. Each hyperplane is uniquely determined by the point at which it intersects the line joining the i^{th} and j^{th} vertex, and the members of the set may be ordered by these points of intersection in exactly the same manner as the breakpoints in the two-box game were. Each such separating hyperplane is transformed into the hyperplane (β_i, β_j) by an optimum sequence.

As an example, let us consider the three-box case where $r_1^3 = r_2^2$. The hypervolume

$$\frac{\beta_j r_i}{\beta_i} \leqslant \frac{p_i}{p_j} \leqslant \frac{\beta_j}{\beta_i r_j}$$

(which contains the recurrent region) is partitioned into $n_1 + n_2 = 5$ hypervolumes where three (n_1) lie to the first side and two (n_2) lie to the second side of (β_i, β_j). This partitioning is illustrated in Fig. 25, where the relative ordering of the looks into boxes 1 and 2 is shown for each hypervolume. Along the line connecting vertex one with vertex two, $p_1 + p_2 = 1$, and the game behaves as though these were the only two boxes involved. Thus, we can determine the positions of each of the separating hyperplanes and the ordering of looks into these two boxes within each hypervolume by using the techniques of Chapter 2. A look into box 1 transforms $\underline{P}$ two (n_2) hypervolumes in the direction of vertex 2, and so forth. If one wishes to find the relative ordering of such looks outside the region enclosed by the hyperplanes $[(\beta_1/r_1), \beta_2]$ and $[\beta_1, (\beta_2/r_2)]$, one can, of course, partition these exterior regions by using the appropriate techniques, which are analogous to those used in the two-box case.

In order to complete the partitioning of the recurrent region, we need only continue the above process for all pairs (i, j). When $r_1^3 = r_2^2 = r_3$ in our three-box example, the general form of the resulting partition is that shown in Fig. 26. In this particular example, we find that five different periodic chains, shown below, occur within the periodic region:



Associated with each of these chains are $\Sigma n_i = 6$ hypervolumes, the sequence of each hypervolume having a different phase. The hypervolumes belonging to each of these chains are indicated in the figure, and it is worth noting that all of those belonging to the same chain have the same general configuration. A more important property to note, however, is that there must always exist more than one periodic chain within the recurrent region when there are more than two boxes. Thus, the state vector $\underline{P}$ will not enter each of the hypervolumes of the recurrent region during one period but will occupy only a subset of them. This contrasts rather strongly with the behavior of the two-box game.

If some of the detection probabilities are equal to one, the recurrent region assumes a new form, and the general behavior of the optimum search sequences within it is also somewhat different. When this occurs, it is convenient to separate the boxes into two sets, letting S' include those which have unity detection probabilities and letting S include the others. Boxes belonging to S' can be examined once, at most, and after each of them has been searched, $\underline{P}$ must belong to the subsimplex generated by the boxes belonging to S. The recurrent region must belong to this subsimplex. It is defined by the relations $p_i = 0$ for all i belonging to S', and

$$\frac{\beta_j r_i}{\beta_i} \leqslant \frac{p_i}{p_j} \leqslant \frac{\beta_j}{\beta_i r_j}$$

for all pairs (i, j) belonging to S. Within this region, the optimum search sequence involves looks into boxes belonging to S and is periodic if there exists a set $\{n_i\}$ such that $r_i^{n_i}$ is the same for all i belonging to S. This recurrent region is partitioned into a set of hypervolumes with unique sequences in the same way as before.

Now that the general behavior of the optimum search strategy has been discussed at some length, it is appropriate to turn to the problem of evaluating the payoff function $U^\infty(P)$. For any fixed search sequence, the payoff is linear in $\underline{P}$, and if we let π_m represent a hypervolume over which a given sequence is optimum, we can express the associated payoff function in the form

$$U_m^\infty(\underline{P}) = \sum_{j=1}^{N} a_m(j)\, p_j \quad .$$

Here, $a_m(j)$ equals the payoff that results if the evader is actually hiding in box j.

If the infinite optimum search sequence associated with π_m is defined, as in Chapter 7, by the set $\{\tau_i(j)\}$, where $\tau_i(j)$ represents the time at which the j^{th} look into box i is completed, each coefficient $a_m(i)$ can be expressed in terms of an infinite series as follows:

$$a_m(i) = q_i \sum_{j=1}^{\infty} r_i^{j-1} \left[\rho_i \left(\frac{1 - d^{\tau_i(j)}}{\alpha} \right) - \eta_i \gamma_i d^{-\tau_i} \sum_{k=1}^{j} d^{\tau_i(k)} - \lambda_i d^{\tau_i(j)} \right] \quad . \tag{8-4}$$

When $\alpha = 0$ (no discounting), this equation reduces to the form

$$a_m(i) = q_i \sum_{j=1}^{\infty} r_i^{j-1} \{ \rho_i \tau_i(j) - j \eta_i \tau_i - \lambda_i \} \quad . \tag{8-5}$$

When a finite sequence transforms π_m into π_n, the payoff $U_m^\infty(P)$ may be expressed as a function of $U_n^\infty(P)$. Letting $\{\tau_i(j)\}$ represent this finite sequence and k_i the total number of looks into box i that are included, we find that

$$a_m(i) = q_i \sum_{j=1}^{k_i} r_i^{j-1} \left[\rho_i \left(\frac{1 - d^{\tau_i(j)}}{\alpha}\right) - \eta_i \gamma_i d^{-\tau_i} \sum_{k=1}^{j} d^{\tau_i(k)} - \lambda_i d^{\tau_i(j)}\right]$$

$$+ r_i^{k_i} \left[\rho_i \left(\frac{1 - d^{\tau_t}}{\alpha}\right) - \eta_i d^{-\tau_i} \sum_{j=1}^{k_i} d^{\tau_i(j)} + d^{\tau_t} a_n(i)\right] \quad , \tag{8-6}$$

and when $\alpha = 0$,

$$a_m(i) = q_i \sum_{j=1}^{k_i} r_i^{j-1} \{\rho_i \tau_i(j) - j\eta_i \tau_i - \lambda_i\}$$

$$+ r_i^{k_i} \{\rho_i \tau_t - k_i \eta_i \tau_i + a_n(i)\} \quad . \tag{8-7}$$

When the search sequence associated with π_m is periodic, the above equations may be used to express each $a_m(i)$ in terms of itself, hence, in closed form.

If the sequence transforming π_m into π_n involves only one look, Eq. (8-6) can be written in the simpler form

$$\pi_m \xrightarrow{i} \pi_n \Longrightarrow$$

$$a_m(i) = (\rho_i - \eta_i)\, \gamma_i - q_i d^{\tau_i} \lambda_i + r_i d^{\tau_i} a_n(i) \quad ,$$

$$a_m(j) = \rho_j \gamma_i + d^{\tau_i} a_n(j) \quad , \quad j \neq i \quad ; \tag{8-8}$$

and Eq. (8-7) reduces to

$$\pi_m \xrightarrow{i} \pi_n \Longrightarrow$$

$$a_m(i) = (\rho_i - \eta_i)\, \tau_i - q_i \lambda_i + r_i a_n(i) \quad ,$$

$$a_m(j) = \rho_j \tau_i + a_n(j) \quad , \quad j \neq i \quad . \tag{8-9}$$

In order to find the evader's optimum strategy in F^∞ and both players' good strategies in G^∞, the space over which $U^\infty(\underline{P})$ is a maximum must be located. Since $U^\infty(\underline{P})$ is linear within each hypervolume π_m over which a single sequence is optimum, the payoff must attain its maximum at at least one extreme point common to a set of such hypervolumes. If this occurs at more than one such point – an unlikely event – these points will generate a space over which $U^\infty(\underline{P})$ is constant. If this space is of degree $N - 1$, it will consist of a single hypervolume π_m. Otherwise, it will form a boundary, of the appropriate degree, that is common to a set of such hypervolumes. The evader's optimum strategy in F^∞ and his good strategy in G^∞ consist in selecting any $\underline{P}$ belonging to this maximum space. The searcher, on the other hand, must find a probability distribution $\underline{Y}_0$ for selecting one of the sequences that are optimum over this space. This distribution must cause the expected payoff to be independent of $\underline{P}$ and hence equal to $\max_{\underline{P}} U^\infty(\underline{P}) \equiv V^\infty$. Such a distribution must exist, since $U^\infty(\underline{P})$ is convex.

Our principal problem is to find a single extreme point at which $U^\infty(\underline{P}) = V^\infty$. The general approach to such a problem is fairly simple. Any extreme point lies at the intersection of at least N – 1 hyperplanes (N – 1 being independent), each of the form (ξ_i, ξ_j) and having the property that the associated optimum sequences will transform it at some point into (β_i, β_j). Radiating from such a point are a number of rays, each formed by the intersection of N – 2 of the independent hyperplanes. If $U^\infty(\underline{P})$ is nonincreasing as $\underline{P}$ moves along each such ray away from the extreme point, this point must be an extreme point of the space over which $U^\infty(\underline{P})$ is a maximum. If $U^\infty(\underline{P})$ is strictly decreasing along each ray, it must be the unique point at which $U^\infty(\underline{P})$ is a maximum. One can start at a known extreme point, $\underline{P}_0$ for example, find the ray along which $U^\infty(\underline{P})$ increases most rapidly, and the next extreme point along this ray. The process can be repeated until an extreme point is found that satisfies the required property. Any pair of extreme points is connected by a network of such rays, and this process will eventually locate the desired point.

Although this process is simple in principle, the computational effort required can quickly reach astronomical proportions as N and Σn_i increase. Radiating from each extreme point are at least 2(N – 1) rays and at an extreme point such as $\underline{P}_0$ there are $2^N - 2$ rays. The task of computing the derivative of $U^\infty(\underline{P})$ along one of these rays is not easy, even after an associated optimum sequence along this ray has been found. Also, the total number of extreme points belonging to the recurrent region alone can be tremendous, even in artificial examples where Σn_i is small. The location of the maximum point within a single hypervolume π_m requires a linear programming routine of no mean size when N is large, and the task of locating a maximum point for the whole simplex can quickly exceed the capabilities of even the largest and fastest computers.

As a result of these considerations, it would be advisable to develop an efficient method for deriving approximately good strategies for the two players. The location of a point reasonably close to the maximum space would suffice as an approximation to the evader's good strategy. Any payoff $U_m^\infty(\underline{P}) = \sum_{j=1}^{N} a_m(j)\, p_j$ that is associated with some optimum sequence has the property that $\min_j \{a_m(j)\} \leqslant V^\infty \leqslant \max_j \{a_m(j)\}$. The sequence limits the evader to $\max_j \{a_m(j)\}$. If the sequence selected is optimum at a point near the maximum space the quantity $\max_j \{a_m(j)\} - \min_j \{a_m(j)\}$ is likely to be small, and with it the searcher should be able to limit the evader to a payoff fairly close to V^∞. In order to get a better solution, the optimum sequences associated with a number of points about the maximum region could be found, and from them a random selection could be made that would yield an expected payoff independent of $\underline{P}$. Such sets do exist as long as there are no inadmissible boxes. Although the resulting payoff will be larger than V^∞, it will be less than the maximum over $\underline{P}$ of any of the individual payoffs.

As shown in Chapter 7, discounting introduces the possibility that some boxes may be inadmissible. If the evader hides in box i and the searcher uses the good strategy that applies when this box is eliminated, the evader will never be found and will receive a payoff equal to ρ_i/α. If, on the other hand, the evader also uses his good strategy that applies when box i is eliminated, the resulting payoff will equal $V^{\infty\prime}$, the value of the reduced game. If $\rho_i/\alpha \leqslant V^{\infty\prime}$, box i is inadmissible, and the good strategies and value of both the original and the reduced game are identical. This condition is both necessary and sufficient. When more than one box is inadmissible, the good strategies and values will be those which apply when all such boxes are removed, and $\rho_i/\alpha \leqslant V^\infty$ for each such box.

Since the above condition depends on the value of the game, or at least on the value of a reduced game that may involve more than one box, there is no simple method for finding the inadmissible boxes when they exist. A stronger but quite simple condition exists, however, that may reveal the presence of such a box in an extreme situation. The value of G^{∞} cannot increase as boxes are removed and must, therefore, be greater or equal to that which applies when only one remains. If both players restrict themselves to box j, the resulting payoff is

$$\frac{(\rho_j - \eta_j)\,\gamma_j - q_j d^{\tau_j} \lambda_j}{1 - r_j d^{\tau_j}} \quad .$$

If there exists a pair of boxes where

$$\frac{\rho_i}{\alpha} \leqslant \frac{(\rho_j - \eta_j)\,\gamma_j - q_j d^{\tau_j} \lambda_j}{1 - r_j d^{\tau_j}} \quad , \tag{8-10}$$

box j dominates box i, and the latter must be inadmissible.

It should be observed that the presence of inadmissible boxes is unlikely unless the interest rate is very high or the earning rates are highly biased. The possibility exists as long as α is unequal to zero, however, and must be taken into account if one wishes to develop a method for finding strategies that approximate the good ones.

While on the subject, it is worthwhile to look ahead and note that the V, the value of game G, is a function of the moving costs. A box may be inadmissible with one set of moving costs but not with another. Since V is monotonically nonincreasing as the moving costs increase, any box inadmissible in G^{∞} will also be inadmissible in G and G°.

8.3 GAME G°

In this section we turn again to the other limiting form of the search evasion game, G°. The N-box form of this game is fortunately quite similar to the two-box form, which we have considered previously, and we also have the good fortune to learn that exact solutions can be found. In order to do this, we must find a state vector $\underline{P}$ that maximizes the evader's guaranteed payoff and another probability vector $\underline{Y}$ with which the searcher can limit the evader to the same amount. The procedures used in the two-box game require little modification. We shall again consider the evader's good strategy first.

As we have mentioned and justified previously in Sec. 3.2, the evader's good strategy in G° must belong to the class of strategies in which the state vector $\underline{P}$ is returned to the same position after each unsuccessful look. If the searcher knows the position of this vector, he may use this information in selecting an optimum search sequence. Since the same $\underline{P}$ applies before each look, a look that is optimum once is always optimum. For a given $\underline{P}$, the searcher can limit the evader to $U^{\circ}(\underline{P}) = \min_i \{U^{\circ}(\underline{P}; i)\}$, where $U^{\circ}(\underline{P}; i)$ is the payoff that results if the searcher always looks into box i. This payoff is

$$U^{\circ}(\underline{P}; i) = \frac{\gamma_i \sum_{j=1}^{N} p_j \rho_j - p_i \left(\gamma_i \eta_i + q_i d^{\tau_i} \lambda_i\right)}{1 - d^{\tau_i}(1 - p_i q_i)} \quad .$$

With no discounting, it reduces to

$$U^\circ(\underline{P}; i) = \frac{\tau_i \sum_{j=1}^{N} p_j \rho_j - p_i(\tau_i \eta_i + q_i \lambda_i)}{p_i q_i} \quad .$$

The evader's good strategy maximizes the guaranteed payoff, and therefore corresponds to that $\underline{P}$ which maximizes $U^\circ(\underline{P})$.

Since each function $U^\circ(\underline{P}; i)$ is nonlinear in $\underline{P}$ for reasons discussed in Sec. 6.3, one cannot be hasty in forming any conclusions regarding the location of the optimum vector. However, the following properties, developed in Appendix D, come to the rescue:

(a) Define $\underline{P}_0$ as a point belonging to the probability simplex that is a solution of the equations

$$U^\circ(\underline{P}; 1) = U^\circ(\underline{P}; 2) = \ldots = U^\circ(\underline{P}; N) \quad .$$

At least one $\underline{P}_0$ must exist, and each one must belong to the interior of the simplex.

(b) All boxes are admissible if and only if there exists a $\underline{P}_0$ that is the unique point in the simplex at which $U^\circ(\underline{P})$ is a maximum. When this occurs, $\underline{P}_0$ must also be the unique point that satisfies the definition in part (a).

(c) If any inadmissible boxes exist, there must be at least one for which

$$\frac{\rho_i}{\alpha} \leqslant U^\circ(\underline{P}_0) \leqslant \max_{\underline{P}} U^\circ(P) \equiv V^\circ \quad .$$

This statement applies for any $\underline{P}_0$.

(d) In the subsimplex generated by the admissible boxes, there exists a unique $\underline{P}$ where $U^\circ(P) = V^\circ$.

The procedure for obtaining the evader's good strategy is, therefore, clear. The set of equations above must first be solved to obtain a trial $\underline{P}_0$. When discounting is not used, this is automatically the solution. With discounting, a check must be made to see whether there are any boxes for which $\rho_i/\alpha \leqslant U^\circ(\underline{P}_0)$. If none exists, the correct solution has been found. If, on the other hand, such boxes do appear, they must be inadmissible and should be eliminated. The above set of equations can then be solved in the reduced game. There is no guarantee that all of the inadmissible boxes, if there are more than one, will be found on the first attempt. The process can be repeated, however, until no more appear. When this occurs, the correct solution has been found. With it, the evader will never hide in any of the inadmissible boxes.

In the preceding list of properties, $\underline{P}_0$ is not claimed to be unique in general because the author was unable to prove that it was true in general. The method for deriving the evader's good strategy does not require this property to be true. It should be stated, however, that the author would be somewhat surprised if an example were found where $\underline{P}_0$ was not unique.

Just as in the two-box form of G°, the searcher's good strategy must belong to that class in which each look is selected according to a probability distribution $\underline{Y} = \{y_i\}$, this distribution being independent of the past search sequence. In a manner analogous to that just used, we can let $W^\circ(\underline{Y}; i)$ represent the payoff that results if the searcher uses $\underline{Y}$ and the evader hides in box i. With such a distribution, the searcher limits the evader to $W^\circ(\underline{Y}) = \max_i W^\circ(\underline{Y}; i)$, and his good strategy is that which minimizes $W^\circ(\underline{Y})$. The expression for $W^\circ(\underline{Y}; i)$ is

$$W^\circ(\underline{Y}; i) = \frac{\rho_i \sum_{j=1}^{N} y_j \gamma_j - y_i \left(\gamma_i \eta_i + q_i d^{\tau_i} \lambda_i\right)}{1 - \sum_{j \neq i} y_j d^{\tau_j} - y_i r_i d^{\tau_i}} ,$$

which reduces to

$$W^\circ(\underline{Y}; i) = \frac{\rho_i \sum_{j=1}^{N} y_j \tau_j - y_i(\tau_i \eta_i + q_i \lambda_i)}{y_i q_i}$$

when discounting is not used.

In contrast to the set $\{U^\circ(P; i)\}$, these payoff functions do not intersect at a common point that satisfies the probability constraints on $\underline{Y}$ if there are any inadmissible boxes. This should cause no difficulty, however, for all the inadmissible boxes, if any, can be found if the evader's good strategy is calculated first. In the reduced game in which all boxes not belonging to S, the set of admissible boxes, have been removed, all of the payoff functions $W^\circ(\underline{Y}; i)$ must intersect at a unique point $\underline{Y}_0$. At this point, where $y_j = 0$ if j does not belong to S, $W^\circ(\underline{Y}_0; i) = W^\circ(\underline{Y}_0)$ for all i belonging to S. In Appendix D it is shown that this point must exist. It is also shown that $W^\circ(\underline{Y}_0) = \min_{\underline{Y}} W^\circ(\underline{Y}) = \max_{\underline{P}} U(P) = V^\circ$. Hence, this $\underline{Y}_0$ is the searcher's good strategy in the original game.

Once the evader's good strategy is known the searcher's can be found more easily, for V° is then known. After removing all of the inadmissible boxes and renumbering the remaining ones from one to N', if necessary, we can write a set of equations, each of the form

$$\frac{\rho_i \sum_{j=1}^{N'} y_j \gamma_j - y_i \left(\gamma_i \eta_i + q_i d^{\tau_i} \lambda_i\right)}{1 - \sum_{j \neq i} y_j d^{\tau_j} - y_i r_i d^{\tau_i}} = V^\circ .$$

These can be rewritten in the form

$$\sum_{j \neq i} y_j \left(\rho_i \gamma_j + V^\circ d^{\tau_j}\right) + y_i \left[\gamma_i(\rho_i - \eta_i) - q_i d^{\tau_i} \lambda_i + r_i d^{\tau_i} V^\circ\right] = V^\circ .$$

Each such equation is linear, and $\underline{Y}_0$ may be obtained by inverting an N'-by-N' matrix.

In the simplified game where α and each η_i and λ_i are equal to zero, the good strategies may be solved algebraically. The solution that results is

$$P_0 = \left\{ \frac{\tau_i / q_i}{\sum_{j=1}^{N} \tau_j / q_j} \right\} , \quad \underline{Y}_0 = \left\{ \frac{\rho_i / q_i}{\sum_{j=1}^{N} \rho_j / q_j} \right\} , \quad V^\circ = \sum_{j=1}^{N} \frac{\tau_i \rho_i}{q_i} .$$

In the more general case, a numerical routine is necessary to find the evader's good strategy. Such a routine should not be too difficult to establish, for the set of functions $\{U^\circ(\underline{P}; i)\}$ is fairly well behaved within the probability simplex. The function $U^\circ(\underline{P}; i)$ is linear over any hyperplane where p_i is fixed and is monotonic along any ray extending from the i^{th} vertex. Also, $U^\circ(\underline{P}; i)$ is equal to a constant over a linear hyperplane.

The good search strategy of G° can be useful when the moving costs are unequal to zero. If the searcher makes each look according to the same probability distribution $\underline{Y}$, the evader does not need to move, even in G°. Rather, he can collect a maximum payoff by remaining in one box. The searcher's good strategy $\underline{Y}_0$ in G° allows the evader to remain in any admissible box. Therefore, $\underline{Y}^\circ$ provides a simple search strategy that limits the evader to V° when the moving costs are unequal to zero.

The good search strategy of G° can also be useful in more practical situations, for it can be used to limit the evader to V° when some of the restrictions imposed by our game model are violated. For example, the evader may not have to hide before the game starts. He may be able to wait until the search process has started and choose a favorable time at which to enter the game. Also, he may be able to stop playing the game temporarily. For example, he may be able to suspend production temporarily while remaining in the same box. Although his earning rate would go to zero, perhaps the detection probability would, also. In some situations it may be cheaper to stop playing than to move, and it may provide a worthwhile evasive device. If the searcher uses his good strategy associated with G°, however, such devices are of no help to the evader. Either the evader should remain in an admissible box for all time or he should not play the game at all. Our search evasion game was motivated by a problem in which the searcher would be very happy if he deterred the evader from playing the game. Naturally, if V° is negative, the evader can receive a larger payoff (zero) if he does not play the game.

Another requirement imposed in our search evasion game was that the moving cost had to be incurred at the time that a move was made. In a more practical situation, a moving cost may result from a decrease in the earning rate over a period of time after the move. It would be extremely difficult to solve a game with this feature. If the good search strategy of G° is used, however, moving can never help the evader. Therefore, this good strategy will limit the evader to V° in this situation also.

8.4 GAME G

When the evader must incur a moving cost whenever he moves, the N-box search evasion game becomes exceedingly complex. In Sec. 8.2, we found that G^∞ could be rather complicated even though its general properties were simple extensions of those of the two-box form. In game G, we are not faced merely with an increase in the size of the problem. Some additional complications arise that do not exist in the two-box game.

In the modified games where the evader reveals the position of the state vector $\underline{P}$ to the searcher before each look, the general approach used in Chapter 4 still holds. Here we must associate a moving cost with each of the possible moves and can let μ_{ij} represent the cost incurred when the evader moves from box i to box j. As before, we can let F represent the game in which the evader can still move before the next look and F' represent that in which this opportunity has passed. The payoff functions that apply in these games when both players use optimum future strategies will be represented by $U(\underline{P})$ and $U'(\underline{P})$.

The functional equation relating $U'(\underline{P})$ to $U(\underline{P})$ is

$$U'(\underline{P}) = \min_i \{U'(\underline{P}; i)\} \quad ,$$

where

$$U'(\underline{P}; i) = \gamma_i \sum_{j=1}^{N} p_j \rho_j - p_i\left(\gamma_i \eta_i + q_i d^{\tau_i} \lambda_i\right) + d^{\tau_i}(1 - p_i q_i)\ U(\underline{P}')$$

and

$$\underline{P} \xrightarrow{i} P' \quad .$$

In F, the evader has the opportunity to move and must weigh the cost of a transformation of the state vector against a possible increase in the future payoff. For a given $\underline{P}$, his optimum strategy has an associated set $\{x_{ij}\}$, where x_{ij} represents the probability that he will move to box j if he is in box i. This produces a transformation to

$$\underline{P}' = \left\{ \sum_{i=1}^{N} p_i x_{ij} \right\} \quad ,$$

and has an associated cost equal to

$$\sum_{i=1}^{N} \sum_{j \neq i} p_i x_{ij} \mu_{ij} \quad .$$

The function $U(\underline{P})$, therefore, must satisfy the functional equation

$$U(\underline{P}) = \max_{\{x_{ij}\}} \left[- \sum_{i=1}^{N} \sum_{j \neq i} (p_i x_{ij} \mu_{ij}) + U'(\underline{P}') \right] \quad ,$$

where

$$\underline{P}' = \left\{ \sum_{i=1}^{N} p_i x_{ij} \right\} \quad .$$

When all the moving costs are identical, the equation reduces to

$$U(\underline{P}) = \max_{\underline{P}'} \left[-\frac{1}{2} \mu \sum_{i=1}^{N} |p_i - p_i'| + U'(\underline{P}') \right]$$

as a result of the efficient move condition.

These functions have properties quite similar to those found in Chapter 4. Both must be continuous and convex. Each will be linear over a set of hypervolumes that may be infinite, and the two are identical over a no-move region. Outside this region, the simplex is partitioned into a set of moving regions within each of which a particular set of moves is required.

Although the solution of these functional equations would be a staggering task, a far more difficult problem arises when we consider the form of the searcher's good strategy. In the two-box form of G, this strategy could be generated by a finite Markov process. Unfortunately, this cannot be done when there are three or more boxes, except in the first strategy interval.

In the two-box game there are only two moving regions, one to each side of the no-move region. Two mixed states and two moving states, at most, are associated with the searcher's

Markov process. When both players use their good strategies, $\underline{P}$ enters a moving region only when the Markov process occupies one of the moving states. The probability distributions associated with the mixed states allow the required move in each state to be admissible.

In the N-box game, on the other hand, the state vector $\underline{P}$ does not simply move back and forth along a line. Instead, it moves in an $N - 1$ space, and can enter many different moving regions. This has rather serious implications, for it is not possible to construct a finite Markov graph that will have a set of mixed and moving states with the required properties. In order to illustrate the problem, we shall examine an extremely simple N-box game for which a partial solution has been found.

8.4.1 Symmetric Three-Box Game with Simple Reward Structure

Let us consider the three-box game where all the detection probabilities are the same and in which the simple reward structure of Chapters 2 through 5 is used. Since this is the simplest game involving more than two boxes, we can be sure that any complications that arise here will arise in general.

With this setup, both G° and G^∞ have trivial solutions. In G°,

$$\underline{P}_0 = \underline{Y}_0 = \{\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\}$$

as a result of symmetry, and $V^\circ = 3/q$. In G^∞, the evader should initially hide according to the probability distribution

$$\underline{P}_0 = \{\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\} \quad ,$$

and the searcher should make an equally likely choice from the 3! periodic search sequences in which the boxes are examined in order. The value is $V^\infty = (3/q) - 1$. It can be shown that $\mu_p = 2$. Hence, when μ exceeds this value, G behaves essentially the same as G^∞.

Over the first strategy interval $0 \leq \mu \leq \mu_1 = 1$, G also has a simple solution. Just as in the two-box game, the evader should return the state vector to the same point after each look. The searcher's good strategy can be generated by a Markov process where each state is defined by the last look. Each such state is both a mixed and a moving state. In such a state, a look into any box is admissible, and the evader may move to the box just searched if he is in any other. These properties hold for the first strategy interval in any game G.

In this example, symmetry causes the solution to be very simple. The evader should always return the state vector to

$$\underline{P}_0 = \{\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\} \quad .$$

The boxes are identical, and letting i, j and k represent the three different boxes, we may write

$$U'(\underline{P}_0) = 1 + \left(\frac{2 + r}{3}\right) U(\underline{P}') \quad ,$$

where $\underline{P}_0 \xrightarrow{i} \underline{P}'$ when box i is examined. In order to return $\underline{P}'$ to $\underline{P}_0$, the evader, if in j or k, should move to box i with probability $q/3$. Therefore,

$$U(\underline{P}') = -\mu \frac{q}{3} \left(\frac{2}{2 + r}\right) + U'(\underline{P}_0) \quad .$$

It follows that $U'(\underline{P}_0) = (3/q) - (2/3)\,\mu = U(\underline{P}_0)$.

In the Markov process that generates the searcher's good strategy, the probability distribution associated with the state σ_i must be of the form $y_i(i) = y$, $y_i(j) = y_i(k) = (1 - y)/2$. State σ_i applies when the last look was into box i. The payoffs associated with σ_i must be of the form

$$W_i(\underline{P}) = W_i'(\underline{P}) = \sum_{j=1}^{N} a_i(j)\, p_j \quad ,$$

where $a_i(i) - a_i(j) = a_i(i) - a_i(k) = \mu$. The solution reveals that $y = (1 - \alpha)/(3 - \alpha)$ and that

$$W_i(\underline{P}) = W_i'(\underline{P}) = \frac{3}{q}\, p_i + \left(\frac{3}{q} - \mu\right) (p_j + p_k) \quad .$$

The searcher should initially make an equally likely look, and he limits the evader to

$$W_0 = \frac{3}{q} - \frac{2}{3}\mu \quad .$$

As long as μ does not exceed one, beyond which y is negative, these strategies are the good strategies, and

$$V(\mu) = \frac{3}{q} - \frac{2}{3}\mu \quad ; \quad 0 \leqslant \mu \leqslant 1 = \mu_1 \quad .$$

Once μ exceeds $\mu_1 = 1$, real problems arise, for the searcher's good strategy may no longer be generated by a simple Markov process. The evader's good strategy is fairly simple, however, and we shall derive it first. The general form of this strategy can be guessed if one considers the behavior of the searcher's good strategy when $\mu = \mu_1$. At this point, $y = 0$ and the searcher never repeats the same look twice in a row. This indicates that the evader should no longer restore $\underline{P}$ to the point at which the payoff is indifferent to any next look. Rather, if box i was examined last, he should restore $\underline{P}$ to a point where

$$p_i = p \quad , \quad p_j = p_k = \frac{1 - p}{2} \quad .$$

If the look preceding the last one was into box j, it is reasonable to assume that the searcher will be more likely to look into box k than into box j. Therefore, let us assume that the only admissible move after looks into boxes j and i, in that order, is from k to i. Before the look into box i, we have

$$p_j = p \quad , \quad p_i = p_k = \frac{1 - p}{2} \quad .$$

After this look,

$$p_j' = \frac{2p}{2p + (1 - p)(1 + r)} \quad , \quad p_i' = \frac{(1 - p)\, r}{2p + (1 - p)(1 + p)} \quad ,$$

and

$$p_k' = \frac{1 - p}{2p + (1 - p)(1 + r)} \quad .$$

The evader must transform this to

$$p_i = p \quad , \quad p_j = p_k = \frac{1 - p}{2}$$

by a possible move from k to i. Thus, p_j' is unaffected, and

$$p_j' = \frac{2p}{2p + (1-p)(1+r)} = \frac{1-p}{2} \quad .$$

Therefore,

$$p = -\frac{(2+r) + \sqrt{5+4r}}{1-r} \quad .$$

The correct transformation occurs if the evader moves from the box that has not been examined during the last two looks (k) to the one just examined (i) with probability

$$x_{ki} = \frac{3 - \sqrt{5+4r}}{2} \quad .$$

At the beginning of the game, the evader should hide at the point

$$\underline{P}_0 = \left\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right\}$$

since $U(\underline{P})$ must be a maximum at this point. If box i is examined first, the evader should move to that box if he is in either of the others with probability $1 - (2+r)\,[(1-p)/2]$, for this strategy transforms $\underline{P}$ to the desired point. Once two or more looks have occurred, the evader moves only in the manner discussed in the preceding paragraph.

With this strategy, the payoff will be independent of where the searcher looks as long as he never examines the same box twice in succession. The guaranteed payoff is

$$U(\underline{P}_0) = \frac{3}{q} - \frac{2}{3}\mu + (\mu - 1)\left[\frac{(3-q)(3-\sqrt{9-4q})}{5q}\right] \quad .$$

The payoff will equal the value as long as the searcher can limit the evader to this amount. Let us first assume that such a search strategy exists (it does) and, in addition, that it is Markovian in form. In the Markov graph, each state would be defined by the last two looks. Thus, we can let σ_{ji} represent the state that applies if the last two looks were made into boxes j and i, where j precedes i. The searcher should not repeat a look into i. Therefore, we can let $y_{ji}(k) = y$ and $y_{ji}(j) = 1 - y$. The payoffs associated with σ_{ji} must be of the form

$$W_{ji}(\underline{P}) = W_{ji}'(\underline{P}) = a_{ji}(i)\, p_i + a_{ji}(j)\, p_j + a_{ji}(k)\, p_k \quad ,$$

where

$$a_{ji}(i) - a_{ji}(k) = \mu \quad \text{and} \quad 0 \leqslant a_{ji}(i) - a_{ji}(j) \leqslant \mu \quad .$$

By means of the usual functional equations, we can find that y does exist. The result is

$$y = \frac{1}{3-\mu}$$

and

$$W_{ji}(\underline{P}) = \frac{3}{q}\, p_i + \left(\frac{3}{q} - 1\right) p_j + \left(\frac{3}{q} - \mu\right) p_k \quad .$$

At the beginning of the game, the searcher has no past search sequence to define a state in the Markov chain. Therefore, we must add the states σ_0, σ_1, σ_2, and σ_3. The state σ_0 is used at the beginning, and one of the others will apply after the first look. Because of symmetry,

$$y_0(i) = \frac{1}{3} \quad \text{for all } i, \quad \text{and} \quad y_i(j) = y_i(k) = \frac{1}{2} \quad .$$

With these probabilities,

$$W_i(\underline{P}) = W_i'(\underline{P}) = \frac{3}{q}\, p_i + [\frac{3}{q} - \frac{1}{2}\, (1 + \mu)]\, (p_j + p_k)$$

and

$$W_0(\underline{P}) = W_0'(\underline{P}) = \frac{3}{q} - \frac{1 + \mu}{3} \quad .$$

Since $W_0(\underline{P}) > U(\underline{P}_0)$, when $\mu > \mu_1$, the strategies we have developed for the two players cannot both be good strategies. The reason for our difficulty becomes apparent if we consider what happens after the first look. If box i is examined first, the evader must move to it, if not already there, with a probability that transforms $\underline{P}$ to $p_i = p$, $p_j = p_k = (1 - p)/2$. On the other hand,

$$W_i(\underline{P}) = \frac{3}{q}\, p_i + [\frac{3}{q} - \frac{1}{2}(1 + \mu)]\, (p_j + p_k) \quad .$$

and

$$[a_i(i) - a_i(j)] = [a_i(i) - a_i(k)] = \frac{1}{2}\,(1 + \mu) < \mu$$

when $\mu > 1$. The searcher's strategy, therefore, causes the necessary moves after the first look to be inadmissible.

These moves must be admissible. A little reflection shows that there can be no satisfactory moving strategy that does not involve a possible move after the first look into the box just examined. Therefore, we must find a search strategy that allows these moves. Such a strategy does exist, in which the searcher introduces a transient into the probability y associated with each state σ_{ji}. This strategy can be found by letting the payoff associated with the state σ_{ji} be a function of the total number of looks that have occurred. Thus, we can let

$$W_{ji}^{n}(\underline{P}) = W_{ji}^{n'}(P) = \sum_{k=1}^{3} a_{ji}^{n}(k)\, p_k \quad .$$

We can also let the associated look probabilities be $y_{ji}^{n}(j)$ and $y_{ji}^{n}(k)$, or y^n and $1 - y^n$. In contrast to the usual recursion equations, a set of difference equations must be written to determine the good strategy. Because of symmetry $[a_{ij}^{n}(k) = a_{ji}^{n}(k) = a_{ki}^{n}(j) \ldots]$, this set can be written in a compact form as follows:

$$a_{ji}^{n}(i) = 1 + a_{ji}^{n+1}(j) \quad ,$$

$$a_{ji}^{n}(j) = 1 + (1 - y^n)\, r a_{ji}^{n+1}(i) + y^n a_{ji}^{n+1}(k) \quad ,$$

$$a_{ji}^{n}(k) = 1 + (1 - y^n)\, a_{ji}^{n+1}(k) + y^n r a_{ji}^{n+1}(i) \quad .$$

Requiring a move from k to j to be admissible adds the equation

$$a_{ji}^{n}(i) = a_{ji}^{n}(k) = \mu \quad .$$

The solution for $a_{ji}^{n}(j)$ is

$$a_{ji}^{n}(j) = \frac{2 + r}{1 - r} + A\left[\frac{1 - \sqrt{5 + 4r}}{2(1 + r)}\right]^{n} + B\left[\frac{1 + \sqrt{5 + 4r}}{2(1 + r)}\right]^{n} .$$

The coefficient B must be set equal to zero so that the transient will decay with time.

The freedom left by the arbitrary coefficient A allows us to make the necessary moves after the first look admissible. That is, we can require that $a_j(j) - a_j(i) = a_j(j) - a_j(k) = \mu$, where $\sigma_j \xrightarrow{i} \sigma^{\circ}_{ji}$. Since σ°_{ji} is entered after two looks have been made, n will be two less than the total number of looks. The solution is

$$a_{ji}^{n}(i) = \frac{3}{q} - (\mu - 1)\left[\frac{-(2 + r) + \sqrt{5 + 4r}}{1 - r^{2}}\right]\left[\frac{1 - \sqrt{5 + 4r}}{2(1 + r)}\right]^{n} ,$$

$$a_{ji}^{n}(j) = \frac{3}{q} - 1 + (\mu - 1)\left[\frac{3 - \sqrt{5 + 4r}}{2(1 - r)}\right]\left[\frac{1 - \sqrt{5 + 4r}}{2(1 + r)}\right]^{n} ,$$

$$a_{ji}^{n}(k) = a_{ji}^{n}(i) - \mu ,$$

$$y^{n} = \frac{1 + (\mu - 1)\left[\frac{1 - \sqrt{5 + 4r}}{2(1 + r)}\right]^{n}}{3 - \mu + (\mu - 1)\left[\frac{3 - \sqrt{5 + 4r}}{2}\right]\left[\frac{1 - \sqrt{5 + 4r}}{2(1 + r)}\right]^{n}} .$$

With these values,

$$W_0(\underline{P}) = \frac{3}{q} - \frac{2}{3}\mu + (\mu - 1)\left[\frac{(3 - q)(3 - \sqrt{9 - 4q})}{6q}\right] .$$

This solution yields the searcher's good strategy in the first part of the second strategy interval. As n approaches infinity, the transients that have been introduced die out, and each term agrees with the corresponding one associated with the simpler search strategy that we tried first. The steady-state value of $y_{ji}^{n}(k) = y^{n}$ is $1/(3 - \mu)$, and this approaches one as μ approaches $\mu_p = 2$.

Unfortunately, the transient introduced into $y_{ji}^{n}(k)$ causes $y_{ji}^{0}(k)$ to equal one when μ is less than μ_p, to be exact, when

$$\mu = \frac{10 + 3r - 4r^{2} + (2 + r)\sqrt{5 + 4r}}{2(4 + 2r - r^{2})} .$$

The first part of the second strategy interval, therefore, ends at this point and does not extend to μ_p.

In this first part, or strategy subinterval, the look probabilities satisfying the difference equations are used after a start-up process of two looks has occurred. The difference equations themselves assure the proper admissible moves once the start-up process has ended. The coefficient A is adjusted to insure that the required moves during the start-up process are admissible.

In succeeding subintervals, the start-up process lasts longer. During this process, some of the looks are made deterministically, and some of the moves that were previously admissible no longer are. Once the start-up process has been completed, the look probabilities are the same as before except that the coefficient A is different. When the start-up process is over, the evader has managed to maneuver $\underline{P}$ into the position where $p_i = p$, $p_j = p_k = (1 - p)/2$, where i represents the last box examined. As μ approaches μ_p, the start-up process becomes infinitely long. The interval $1 \leqslant \mu \leqslant 2 = \mu_p$ is, therefore, partitioned into an infinite number of strategy subintervals. The start-up process differs from subinterval to subinterval, whereas the general behavior thereafter is the same for all those subintervals belonging to the same strategy interval (μ_1, μ_p). In a more general N-box game, there may be a finite number of strategy intervals, but there will always be an infinite number of subintervals in all but the first.

8.4.2 Approximate Solutions for G

In the example just discussed, an exact solution of G was found in the first strategy interval and in the first subinterval of the second. Finding the solution in any other of these subintervals would be an enormous task, and it should be apparent that the prospects of finding an exact solution to a more general N-box form of G are slight, if not out of the question.

Methods for finding approximate solutions are, therefore, in order. Although no method has been developed, a general approach to finding an approximately good search strategy is suggested for future research. With this approach, we assume that the searcher has a poor memory and can remember only where he has made his last n looks. Since he can use only information he remembers, his optimum search strategy under this condition can be generated by a Markov process. In particular, each recurrent state in the process is defined by where the last n looks were made. Transient states that apply when fewer than n looks have been made also exist.

As an example, let us consider the approach that could be taken when n = 2. For simplicity, we shall let $\rho_i = \tau_i = 1$ and $\eta_i = \lambda_i = 0$. Letting σ_{ji} represent the state that applies when the last look was made into box i and was preceded by a look into j, we can express the payoffs associated with σ_{ji} in the form

$$W'_{ji}(\underline{P}) = \sum_{k=1}^{N} a'_{ji}(k)\, p_k$$

and

$$W_{ji}(\underline{P}) = \sum_{k=1}^{N} a_{ji}(k)\, p_k \quad .$$

The payoff $W_{ji}(\underline{P})$ applies when the evader can move before the next look and $W'_{ji}(\underline{P})$ applies when this opportunity has been lost. The cost of moving from box k to box ℓ is $\mu_{k\ell}$. Therefore,

$$a_{ji}(k) = \max_{\ell} \{-\mu_{k\ell} + a'_{ji}(\ell)\} \quad ,$$

where $\mu_{kk} = 0$.

With each state σ_{ji} we can associate a set of look probabilities $y_{ji}(k)$ where $\sum_{k=1}^{N} y_{ji}(k) = 1$. If box k is examined, a transition to σ_{ik} occurs. Therefore,

$$a'_{ji}(k) = 1 + \sum_{\ell \neq k} y_{ji}(\ell)\, a_{i\ell}(k) + y_{ji}(k)\, r_k a_{ik}(k) \quad .$$

Similar equations can be written for each of the transient states. At the beginning of the game, the payoff $W_0(\underline{P}) = \sum_{k=1}^{N} a_0(k)\, p_k$ applies, and the searcher limits the evader to $\max_k \{a_0(k)\} = \max_k \{a_0'(k)\}$. To find the look probabilities that minimize the above expression, a nonlinear programming routine is necessary. It is this routine that must be studied in detail.

As n approaches infinity, the searcher's memory improves, and in the limit, the approximating strategy approaches the good search strategy. In the process, the number of states in the Markov process can increase rapidly. With a memory of n looks, there can be N^n recurrent states and $\sum_{j=0}^{n-1} N^j$ transient states.

The total number may not be as large, however, for some states may be superfluous. When $n = 3$, for example, $y_{ijk}(\ell)$ may equal zero for all i. If this occurs, the state $\sigma_{jk\ell}$ will never be entered and is of no interest. The problem, of course, is to find a method for predicting such an event before the solution is attempted. Such a method may lie in solving the $n - 1$ approximation first. It appears likely that if $y_{jk}(\ell) = 0$ for all j in the $n - 1$ approximation, $y_{ijk}(\ell)$ must also equal zero for all i, j in the n approximation. If this property can be proved to exist, the problem of finding a good approximation to the searcher's good strategy can be greatly simplified. The solutions to the approximations of order 0, 1, 2, . . . can be found in order, and the process can be stopped when diminishing returns are found or when the computational effort becomes too large.

CHAPTER 9
CONCLUSION

The game that we have considered is a two-sided extension of a one-sided search problem. Although all search problems need not be considered from a two-sided point of view, this is sometimes necessary. In our game, the search is directed against a conscious evader or an object controlled by such an evader. The evader can observe the searcher's actions and can capitalize on any errors he makes. At the beginning of the game, the evader hides in one of several boxes, each of which has an associated detection probability. The search process consists of a sequence of looks into the various boxes until the evader is found. Each look into a given box takes a fixed amount of time. A particular evasion device – moving between looks – has been assumed. The game is zero-sum, and a fairly general reward structure that can include discounting has been developed. The reward coefficients associated with this structure, as well as the location of the boxes and their detection probabilities, are known to both players.

We have been able to derive the good strategies for the two players when the game involves two boxes. In G^∞, exact solutions can be obtained when there exists a pair of integers n_1 and n_2 such that $r_1^{n_1} = r_2^{n_2}$. The escape probabilities r_1 and r_2 are the complements of the detection probabilities. An exact solution can also be found if one or both of the detection probabilities are equal to unity. When the ratio of the escape probabilities is irrational, an approximate solution can be obtained. This approximation can be made to any desired degree of accuracy. In game G, where moving is allowed at a cost, the solution to the searcher's good strategy is identical to that of G^∞ if the moving costs are prohibitive. When these moving costs are not prohibitive, the exact good strategies can be found in general. Exact solutions can also be obtained in G°, where the moving costs are equal to zero.

The search evasion game becomes much more complex when there are three or more boxes. Although G° may still be solved exactly, the computational effort required to solve G^∞ can be prohibitive. When more than two boxes are involved, the general properties of G become quite different, and the good search strategy can no longer be generated by a finite Markov process. The limited memory approach to finding an approximation to the good search strategy is suggested for future research. Such a strategy can still be generated by a Markov process.

The results of our study of F^∞ may be useful in treating some one-sided search problems. The optimum search strategy is quite simple when the position of the evader, or object, can be described by a probability vector. Only the problem of locating the point at which $U^\infty(P)$ is a maximum causes the solution of the N-box form of G^∞ to be difficult. If the object is not a conscious entity whose motives are opposed to those of the searcher, we have no reason to suppose that the worst of all probability vectors applies. Any reasonable statistical estimate of the position of the object should be preferable to taking the pessimist's approach, i.e., using the minimax solution.

Most of the reward structure that has been developed could be useful in treating a one-sided search problem of this type. The detection loss λ_i could be used to represent a reward associated with finding the object. The reward associated with a given look could be used to represent the cost of making the look. It would be difficult to imagine a problem in which a look was less costly to make if the object were in the box examined. Therefore, the η's would normally be equal to zero. In most situations, there would also be no reason to suppose that the earning rates varied from box to box. If one wished to locate a faulty part that was causing damage over time in a complex system, however, they might be of use. Clearly, the set of look times and discounting could be useful in a practical one-sided search problem.

At this point, it is worthwhile to review some of the qualitative aspects of the good strategies associated with the two-box form of G. Let us first consider the evader's good strategy. If an arbitrary strategy is assigned to the evader, we may define his position as the search process proceeds by means of a probability vector. If the probability that he is in one box becomes sufficiently large, the evader should move from this box if he is there with a certain probability. This causes the probability vector describing his position to be transformed to the nearest boundary of the no-move region.

The searcher's good strategy can be generated by a finite Markov process. In some states of this process, the next look is made deterministically. In others, the mixed states, the next look is made according to a probability distribution. When the searcher uses his good strategy, the evader will collect a payoff equal to the value if he never moves. That is, not moving is always an admissible alternative of the evader's good strategy. In certain situations, a particular move is also admissible. As the moving costs increase, deterministic looks are made more frequently, and the situations in which a move is admissible occur less frequently.

When the moving costs are prohibitive, the searcher's good strategy is identical to the one that applies in G^{∞}, the game in which moving is prohibited. In this strategy, the searcher makes a random selection from two infinite ~~search~~ sequences. Once this choice has been made, the search process is completely deterministic. This strategy minimizes the payoff that results if the evader never moves. The evader should not move because such an action can only decrease the payoff.

When the moving costs are not prohibitive, the searcher should not use his good strategy that applies in G^{∞}. If he were to use this strategy, the evader could gain a definite advantage (perhaps a very large one) by using a strategy that involved some deterministic moves. No search strategy that allows the evader to gain a definite advantage by moving can be the good search strategy. Therefore, the good search strategy is the one that minimizes the no-move payoff without violating this condition.

In a sense, the good search strategy maximizes the number of situations in which moving is an admissible alternative subject to the above condition. In each strategy interval, a particular transition diagram is associated with the Markov process that generates the good search strategy. This diagram includes one or two moving states. In such a state, a particular move, as well as not moving, is an admissible alternative. In the remaining states, no moves are admissible. If the moving costs are increased sufficiently, a new strategy interval will apply. In the associated transition diagram, there are more states in which no moves are admissible. One cannot use the previous transition diagram in this strategy interval. If one were to try, he would find that some of the transition probabilities, or look probabilities, associated with the mixed states would be negative.

In the N-box form of G, the good search strategy cannot be generated by a finite Markov process. Associated with the evader's good strategy is a set of situations in which moving is required with a nonzero probability. No Markovian search strategy can yield a payoff that is indifferent to whether or not the move is made in each of these situations. (The one exception to this statement applies in the first strategy interval.) In the symmetric three-box example that was solved, we saw that there was a Markovian search strategy which allowed all of the moves associated with the evader's good strategy to be admissible except those that applied during the start-up process. Unfortunately, the start-up process occurs at the beginning of the game, and the early behavior of the game has the strongest influence on the payoff.

In G°, no cost is incurred by the evader when he moves. As a result, the searcher cannot gain any inference concerning the evader's position from his past sequence of unsuccessful looks, and each look should be made according to the same probability distribution. The good search strategy causes the payoff to be independent of the evader's position and vitiates the influence of moving (all moves are admissible).

When the N-box form of G° was considered, we saw that the associated good search strategy may be useful when evasion devices not included in our game are considered. For example, the evader may be able to select a favorable time after the search has started to enter the game. He may also be able to temporarily suspend production or leave the game. These additional devices do not aid the evader if the searcher uses his good strategy associated with G°. This strategy would not be the good strategy in the more general game of this type. It can be calculated, however, and it may prove useful in a practical two-sided search problem involving these additional devices.

Although the search evasion game we have studied includes only one evasion device, it has demonstrated the interesting influence that a conscious evader can have on the outcome of a search process. This study should only whet the appetite for deeper studies. In a more general game, it will be more difficult to find exact solutions. In fact, there is no reason to suppose that a value and good strategies will always exist. Nevertheless, the further development of relatively efficient search procedures that take the actions of the evader into account should prove interesting and useful.

REFERENCES

1. B.O. Koopman, "The Theory of Search, Part I, Kinematic Basis," Operations Research 4, 324 (1956).
2. ____________, "The Theory of Search, Part II, Target Detection," Operations Research 4, 503 (1956).
3. ____________, "The Theory of Search, Part III, The Optimum Distribution of Search Effort," Operations Research 5, 613 (1957).
4. B. Gluss, "An Optimum Policy for Detecting a Fault in a Complex System," Operations Research 7, 468 (1959).
5. S.M. Pollock, "Optimal Sequential Strategies for Two-Region Search when Effort is Quantized," Interim Technical Report No. 14, M.I.T. Project, Fundamental Investigations in Methods of Operations Research.
6. N.M. Blackman, "Prolegomena to Optimum Discrete Search Procedures," Naval Research Logistics Quart. 6, 273 (1959).
7. ____________, "Optimum Search for Objects Having Unknown Arrival Times," Operations Research 7, 625 (1959,.
8. L.E. Dubbins, "A Discrete Evasion Game," in Contributions to the Theory of Games (Princeton University Press, Princeton, New Jersey, 1957), Vol. III, pp. 231-255.
9. R. Luce and H. Raiffa, Games and Decisions (Wiley, New York, 1957).

APPENDIX A
THE SEARCHER'S OPTIMUM STRATEGY IN F^{∞}

The searcher's optimum strategy in F^{∞} requires each next look to be made into a box for which $p_i\beta_i$ is a maximum, where $\underline{P} = \{p_i\}$ represents the value of the probability vector that applies when the decision is made and

$$\beta_i = \frac{q_i d^{\tau_i}}{\gamma_i} (\rho_i + \alpha\lambda_i) + \alpha\eta_i \quad .$$

In order to prove this, let us adopt the following notation. Let S represent an infinite search sequence, and let $S = (s_i, s_j, \ldots, S_k)$ represent a partition of this sequence into an ordered set of subsequences where $s_i, s_j, \ldots$ are finite and S_k is infinite. Let τ_i represent the length of s_i in time. Let $\underline{P}_{i,j,\ldots}$ represent the a posteriori position into which the a priori vector $\underline{P}$ is transformed by the sequence $(s_i, s_j, \ldots)$ if detection does not occur.

Letting $U^{\infty}(\underline{P}; S) = U^{\infty}(\underline{P}; s_i, s_j, S_k)$ represent the payoff given $\underline{P}$ and $S = (s_i, s_j, S_k)$, we may express this payoff in the form

$$U^{\infty}(\underline{P}; s_i, s_j, S_k) = f(\underline{P}; s_i) + g(\underline{P}; s_i)\, d^{\tau_i}\, U^{\infty}(\underline{P}_i; s_j, S_k)$$

$$= f(\underline{P}; s_i, s_j) + g(\underline{P}; s_i, s_j)\, d^{\tau_i + \tau_j}\, U^{\infty}(\underline{P}_{i,j}; S_k) \quad .$$

The contribution $f(\underline{P}; s_i)$ to the payoff occurs during the subsequence s_i when $\underline{P}$ is the a priori value of the state vector; $g(\underline{P}; s_i)$ equals the probability that detection does not occur during s_i.

Let us consider an arbitrary infinite sequence S_I and an arbitrary a priori $\underline{P}$. Let s_a represent the subsequence of S_I that starts at the beginning of S_I and continues up to, but does not include, the first look that violates the optimum search rule (s_a may be empty). Let C represent the set of boxes that could be examined on the next look without violating the optimum search rule, and let s_b be the maximum subsequence following s_a that does not include a look into a box belonging to C. Letting B represent the set of boxes that are examined at least once in s_b, we see that $B \cap C = \phi$.

The next look following s_b is into a box belonging to C. This box will be called box c and the look into this box will form the subsequence s_c. Let S_d be the remainder of S_I. For the moment, let us assume that s_b is finite. Then $S_I = (s_a, s_b, s_c, S_d)$.

By reversing the order of s_b and s_c, we can define a new sequence

$$S_{II} = (s_a, s_c, s_b, S_d) \quad .$$

This definition will extend the sequence of optimum looks by one look $[s'_a = (s_a, s_c)]$, and we must show that

$$\Delta = U^{\infty}(\underline{P}; S_I) - U^{\infty}(\underline{P}; S_{II}) > 0 \quad .$$

These payoffs may be expressed in the form

$$U^{\infty}(\underline{P}; S_I) = f(\underline{P}; s_a) + g(\underline{P}; s_a)\, d^{\tau_a}\, U^{\infty}(\underline{P}_a; s_b, s_c, S_d) \quad ,$$

$$U^{\infty}(\underline{P}; S_{II}) = f(\underline{P}; s_a) + g(\underline{P}; s_a)\, d^{\tau_a}\, U^{\infty}(\underline{P}_a; s_c, s_b, S_d) \quad .$$

Therefore, Δ is strictly greater than zero if and only if $U^{\infty}(\underline{P}_a; s_b, s_c, S_d)$ is strictly greater than $U^{\infty}(\underline{P}_a; s_c, s_b, S_d)$. Furthermore,

$$U^{\infty}(\underline{P}_a; s_b, s_c, S_d) = f(\underline{P}_a; s_b) + g(\underline{P}_a; s_b)\, d^{\tau_b} f(\underline{P}_{a,b}; s_c)$$

$$+ g(\underline{P}_a; s_b, s_c)\, d^{\tau_b+\tau_c}\, U^{\infty}(\underline{P}_{a,b,c}; S_d)$$

and

$$U^{\infty}(\underline{P}_a; s_c, s_b, S_d) = f(\underline{P}_a; s_c) + g(\underline{P}_a; s_c)\, d^{\tau_c} f(\underline{P}_{a,c}; s_b)$$

$$+ g(\underline{P}_a; s_c, s_b)\, d^{\tau_c+\tau_b}\, U^{\infty}(\underline{P}_{a,c,b}; S_d) \quad .$$

But, $g(\underline{P}_a; s_b, s_c) = g(\underline{P}_a; s_c, s_b)$ and $\underline{P}_{a,b,c} = \underline{P}_{a,c,b}$. Therefore,

$$\Delta = f(\underline{P}_a; s_b) + g(\underline{P}_a; s_b)\, d^{\tau_b} f(\underline{P}_{a,b}; s_c)$$

$$- f(\underline{P}_a; s_c) + g(\underline{P}_a; s_c)\, d^{\tau_c} f(\underline{P}_{a,c}; s_b) \quad .$$

We shall now prove that $\Delta > 0$. For convenience, let us simplify our notation by using $\underline{P}$ in place of $\underline{P}_a$. Then,

$$\text{for all} \quad i \in C, \text{ and all} \quad j \notin C: \quad p_i\beta_i > p_j\beta_j \quad ;$$

$$\text{for all} \quad i \in C, \text{ and all} \quad j \in B: \quad p_i\beta_i > p_j\beta_j \quad .$$

For each $i \in B$, let $\tau_i(n)$ equal the time in s_b at which the n^{th} look into box i is completed. Also, let k_i represent the total number of looks into box i contained in s_b. Then

$$f(\underline{P}; s_b) = \sum_{i \in B} \left\{ p_i q_i \sum_{j=1}^{k_i} r_i^{j-1} \left[\rho_i \left(\frac{1 - d^{\tau_i(j)}}{\alpha} \right) - \eta_i \gamma_i d^{-\tau_i} \sum_{k=1}^{j} d^{\tau_i(k)} - \lambda_i d^{\tau_i(j)} \right] \right.$$

$$\left. + p_i r_i^{k_i} \left[\rho_i \left(\frac{1 - d^{\tau_b}}{\alpha} \right) - \eta_i \gamma_i d^{-\tau_i} \sum_{j=1}^{k_i} d^{\tau_i(j)} \right] \right\}$$

$$+ p_c \rho_c \left(\frac{1 - d^{\tau_b}}{\alpha} \right) + \sum_{\substack{i \neq c, \\ \notin B}} p_i \rho_i \left(\frac{1 - d^{\tau_b}}{\alpha} \right) \quad ;$$

$$g(\underline{P}; s_c)\, d^{\tau_c} f(\underline{P}_c; s_b) = d^{\tau_c} \sum_{i \in B} \left\{ p_i q_i \sum_{j=1}^{k_i} r_i^{j-1} \left[\rho_i \left(\frac{1 - d^{\tau_i(j)}}{\alpha} \right) - \eta_i \gamma_i d^{-\tau_i} \sum_{k=1}^{j} d^{\tau_i(k)} \right.\right.$$

$$\left. - \lambda_i d^{\tau_i(j)} \right] + p_i r_i^{k_i} \left.\left[\rho_i \left(\frac{1 - d^{\tau_b}}{\alpha} \right) - \eta_i \gamma_i d^{-\tau_i} \sum_{j=1}^{k_i} d^{\tau_i(j)} \right]\right\}$$

$$+ p_c d^{\tau_c} r_c \rho_c \left(\frac{1 - d^{\tau_b}}{\alpha} \right) + d^{\tau_c} \sum_{\substack{i \neq c, \\ \notin B}} p_i \rho_i \left(\frac{1 - d^{\tau_b}}{\alpha} \right) \quad ;$$

$$f(\underline{P}; s_c) = \sum_{i \in B} p_i \rho_i \gamma_c + p_c \left[\gamma_c (\rho_c - \eta_c) - q_c d^{\tau_c} \lambda_c \right] + \sum_{\substack{i \neq c \\ i \notin B}} p_i p_i \gamma_c \quad ;$$

$$g(P; s_b)\, d^{\tau_b} f(P_b; s_c) = d^{\tau_b} \sum_{i \in B} p_i r_i^{k_i} \rho_i \gamma_c + d^{\tau_b} p_c \left[\gamma_c (\rho_c - \eta_c) - q_c d^{\tau_c} \lambda_c \right]$$

$$+ d^{\tau_b} \sum_{\substack{i \neq c \\ \notin B}} p_i \rho_i \gamma_c \quad .$$

By collecting terms and noting that $\gamma_c = (1 - d^{\tau_c})/\alpha$ and $\gamma_i = (1 - d^{\tau_i})/\alpha$, we can reduce Δ to the form

$$\Delta = - \sum_{i \in B} p_i \rho_i \left(\frac{1 - d^{\tau_c}}{\alpha} \right) q_i \sum_{j=1}^{k_i} r_i^{j-1} d^{\tau_i(j)}$$

$$- \sum_{i \in B} p_i \eta_i \frac{(1 - d^{\tau_c})(1 - d^{\tau_i})}{\alpha} d^{-\tau_i} \left[q_i \sum_{j=1}^{k_i} r_i^{j-1} \sum_{k=1}^{j} d^{\tau_i(k)} + r_i^{k_i} \sum_{j=1}^{k_i} d^{\tau_i(j)} \right]$$

$$- \sum_{i \in B} p_i \lambda_i (1 - d^{\tau_c})\, q_i \sum_{j=1}^{k_i} r_i^{j-1} d^{\tau_i(j)} + p_c \rho_c q_c d^{\tau_c} \left(\frac{1 - d^{\tau_b}}{\alpha} \right)$$

$$+ p_c \eta_c \frac{(1 - d^{\tau_c})(1 - d^{\tau_b})}{\alpha} + p_c q_c \lambda_c d^{\tau_c} (1 - d^{\tau_b}) \quad .$$

But,

$$\sum_{j=1}^{k_i} r_i^{j-1} d^{\tau_i(j)} \leq \sum_{j=1}^{k_i} d^{\tau_i(j)}$$

and

$$q_i \sum_{j=1}^{k_i} r_i^{j-1} \sum_{k=1}^{j} d^{\tau_i(k)} + r_i^{k_i} \sum_{j=1}^{k_i} d^{\tau_i(j)} \leqslant \sum_{j=1}^{k_i} d^{\tau_i(j)} \quad .$$

Therefore,

$$\Delta \geqslant \Delta_1 = - \sum_{i\epsilon B} p_i(1 - d^{\tau_c}) \left[\sum_{j=1}^{k_i} d^{\tau_i(j)}\right] \left[\frac{\rho_i q_i}{\alpha} + \left(\frac{1 - d^{\tau_i}}{\alpha}\right) d^{-\tau_i}\eta_i + q_i\lambda_i\right]$$

$$+ p_c [1 - d^{\tau_b}] \left[\frac{\rho_c q_c}{\alpha} d^{\tau_c} + \left(\frac{1 - d^{\tau_c}}{\alpha}\right) \eta_c + q_c\lambda_c d^{\tau_c}\right] \quad .$$

But,

$$1 - d^{\tau_b} = \sum_{i\epsilon B} d^{-\tau_i}(1 - d^{\tau_i}) \sum_{j=1}^{k_i} d^{\tau_i(j)} \quad ;$$

therefore,

$$\Delta \geqslant \Delta_1 = \sum_{i\epsilon B} \gamma_i\gamma_c d^{-\tau_i} \left[\sum_{j=1}^{k_i} d^{\tau_i(j)}\right] (p_c\beta_c - p_i\beta_i) \quad .$$

Since $p_c\beta_c > p_i\beta_i$ for all i belonging to B,

$$\Delta \geqslant \Delta_1 > 0 \quad .$$

If the sequences s_b and s_c are reversed, therefore, the payoff is reduced. If s_b were to increase in length and become infinite, the inequality $\Delta \geqslant \Delta_1$ could only become stronger, since some of the members of $\{k_i\}$ must eventually become infinite. Therefore, if $S_I = (s_a, S_b)$, the payoff can be reduced by inserting a look into a box belonging to C between s_a and S_b. Hence, the sequence $S_{II} = (s_a, s_c, S_b)$ would yield a lower payoff than $S_I = (s_a, S_b)$. The process can be continued indefinitely until S_I has been replaced by a sequence S_a. This yields the minimum payoff, and the assumed optimum search rule is indeed optimum.

APPENDIX B

SOME PROPERTIES OF THE TWO-BOX MODIFIED GAMES F AND F'

Consider the sets of truncated modified games $\{F_n\}$ and $\{F'_n\}$. These games have the same definitions as F and F' except that F'_n follows F_n and F_{n-1} follows F'_n. In F_0, play stops and the evader collects the payoff V°, the value of G°. Associated with F_n and F'_n are the payoff functions $U_n(P)$ and $U'_n(P)$, respectively, which apply when both players use optimum strategies. The functional equations are

$$U'_n(P) = \min \begin{cases} U'_n(P;1) = P\left[\gamma_1(\rho_1 - \eta_1) - q_1 d^{\tau_1}\lambda_1\right] \\ \qquad + (1-P)\,\gamma_1\rho_2 \\ \qquad + [Pr_1 + 1 - P]\; d^{\tau_1} U_{n-1}\left[\dfrac{Pr_1}{Pr_1 + 1 - P}\right] \\ U'_n(P;2) = P\,\gamma_2\rho_1 \\ \qquad + (1-P)\left[\gamma_2(\rho_2 - \eta_2) - q_2 d^{\tau_2}\lambda_2\right] \\ \qquad + [P + (1-P)\,r_2]\; U_{n-1}\left[\dfrac{P}{P + (1-P)\,r_2}\right] \end{cases}$$

and

$$U_n(P) = \max_{P'} \begin{cases} -\mu_1(P' - P) + U'_n(P') \quad , & P' \geqslant P \\ -\mu_2(P - P') + U'_n(P') \quad , & P' \leqslant P \end{cases} \quad ,$$

where

$$U_0(P) = V° \quad .$$

We shall require both boxes to be strictly admissible; that is,

$$\frac{\rho_1}{\alpha} > \frac{(\rho_2 - \eta_2)\,\gamma_2 - q_2 d^{\tau_2}\lambda_2}{1 - r_2 d^{\tau_2}} \quad ,$$

$$\frac{\rho_2}{\alpha} > \frac{(\rho_1 - \eta_1)\,\gamma_1 - q_1 d^{\tau_1}\lambda_1}{1 - r_1 d^{\tau_1}} \quad .$$

Theorem 1.

For all $n > 0$: $U_n(P)$ and $U'_n(P)$ are continuous and convex.

In F_0, the function $U_0(P) = V°$. Hence, $U_0(P)$ is continuous and convex. To prove the theorem, we shall show that

$U_{n-1}(P)$ is continuous and convex $\Longrightarrow$ $U_n(P)$ and $U'_n(P)$ are continuous and convex.

For convenience, we shall also use the same technique to show that $U_n(P)$ and $U_n'(P)$ are both piecewise linear for all finite n.

Assume that $U_{n-1}(P)$ is continuous, piecewise linear, and convex. Consider

$$U_n'(P;1) = P\left[\gamma_1(\rho_1 - \eta_1) - q_1 d^{\tau_1} \lambda_1\right]$$
$$+ (1-P)\,\gamma_1\rho_2$$
$$+ (Pr_1 + 1 - P)\, U_{n-1}\left[\frac{Pr_1}{Pr_1 + 1 - P}\right] \quad .$$

Clearly, $U_n'(P;1)$ must be continuous since $U_{n-1}(P)$ is. The function $U_{n-1}(P)$ is linear over a set of intervals $\{\pi_i^{n-1}\}$ that form a partition of the interval (0, 1). Over π_i^{n-1}, we may express $U_{n-1}(P)$ in the form

$$U_{n-1}(P) = a_i^{n-1}\, P + b_i^{n-1}\,(1-P) \quad .$$

Define π_i^n by the relation

$$\pi_i^n \xrightarrow{1} \pi_i^{n-1} \quad .$$

For all $P \in \pi_i^n$:

$$U_n'(P;1) = P\left[\gamma_1(\rho_1 - \eta_1) - q_1 d^{\tau_1}\lambda_1 + r_1 d^{\tau_1} a_i^{n-1}\right]$$
$$+ (1-P)\left[\gamma_1\rho_2 + d^{\tau_1} b_i^{n-1}\right]$$
$$= a_i^n P + b_i^n (1-P) \quad .$$

Hence, $U_n'(P;1)$ is piecewise linear over each interval belonging to $\{\pi_i^n\}$, where $\{\pi_i^n\}$ partitions the interval (0, 1).

Let $\pi_j^{n-1} > \pi_i^{n-1} \Longrightarrow$ for all $P_i \in \pi_i^{n-1}$, and for all $P_j \in \pi_j^{n-1}$: $P_j \geqslant P_i$. The function $U_{n-1}(P)$ is convex. Hence,

$$\pi_j^{n-1} > \pi_i^{n-1} \Longleftrightarrow a_j^{n-1} < a_i^{n-1} \quad \text{and} \quad b_j^{n-1} > b_i^{n-1} \quad .$$

But,

$$\pi_j^{n-1} > \pi_i^{n-1} \Longleftrightarrow \pi_j^n > \pi_i^n \quad ,$$
$$a_j^{n-1} < a_i^{n-1} \Longleftrightarrow a_j^n < a_i^n \quad ,$$
$$b_j^{n-1} > b_i^{n-1} \Longleftrightarrow b_j^n > b_i^n \quad .$$

Therefore,

$$\pi_j^n > \pi_i^n \Longleftrightarrow a_j^n < a_i^n \quad \text{and} \quad b_j^n > b_i^n \quad ,$$

and $U_n'(P;1)$ is also convex.

The same reasoning can be used to show that $U_n'(P;2)$ is continuous, piecewise linear, and convex:

$$U_n'(P) = \min \begin{cases} U_n'(P;1) \\ U_n'(P;2) \end{cases} \quad .$$

Therefore, $U_n'(P)$ is also continuous and convex. It is piecewise linear as long as $U_n'(P;1)$ and $U_n'(P;2)$ do not intersect at an infinite number of points. In proving Theorem 2, we shall show that these functions intersect at a unique point $P_0^n \in (0,1)$. Hence, $U_n'(P)$ is piecewise linear.

$U_n(P)$ can be constructed from $U_n'(P)$ by using the techniques discussed in Sec. 4.3, and $U_n(P)$ must also be continuous, piecewise linear, and convex.

Theorem 2.

In F_n', there exists a unique P_0^n:

$$0 < P_0^n < 1 \quad ,$$

$$P < P_0^n \Longrightarrow U_n'(P;1) > U_n'(P;2) \quad ,$$

$$P > P_0^n \Longrightarrow U_n'(P;1) < U_n'(P;2) \quad .$$

Let us adopt the notation

$$\delta_n'(P';i)+ \equiv \left. \frac{dU_n'(P;i)}{dP} \right|_{P'+}$$

$$\delta_n'(P';i)- \equiv \left. \frac{dU_n'(P;i)}{dP} \right|_{P'-} \quad .$$

To ease the notation further, we shall assume that a statement concerning $\delta_n'(P;i)+$ over an interval (P_j, P_k) applies only for $P_j \leqslant P < P_k$, unless an explicit statement is made to the contrary. Similarly, a statement concerning $\delta_n'(P;i)-$ over (P_j, P_k) will apply only if $P_j < P \leqslant P_k$. A statement concerning an unsigned quantity $\delta_n'(P;i)$ will apply to both $\delta_n'(P;i)+$ and $\delta_n'(P;i)-$ once the above condition is imposed. If a statement concerns several unsigned quantities, such as $\delta_n'(P;1)$ and $\delta_n'(P;2)$, it will be inferred that the statement applies as long as both quantities are evaluated by taking the limit of the derivative in the same way.

We shall prove the theorem by demonstrating a stronger property; namely,

for all $P \in (0,1)$:

$$\delta_n'(P;2) > \delta_n'(P;1)$$

and

$$U_n'(0;2) < U_n'(0;1)$$

$$U'_n(1;2) > U'_n(1;1) \quad ,$$

where $n \geqslant 1$.

First consider F'_1:

$$U'_1(P;1) = P\left[\gamma_1(\rho_1 - \eta_1) - q_1 d^{\tau_1}\lambda_1\right] + (1-P)\,\gamma_1\rho_2$$

$$+ [Pr_1 + 1 - P]\, d^{\tau_1} V^\circ \quad ,$$

$$U'_1(P;2) = P\gamma_2\rho_1 + (1-P)\left[\gamma_2(\rho_2 - \eta_2) - q_2 d^{\tau_2}\lambda_2\right]$$

$$+ [P + (1-P)\,r_2]\, d^{\tau_2} V^\circ \quad .$$

The game F'_1 is equivalent to F° when the evader is not allowed to move until after the first look. Both boxes are strictly admissible. The functions $U'_1(P;1)$ and $U'_1(P;2)$ must intersect at the point P_0 that corresponds to the evader's good strategy in G°, where $0 < P_0 < 1$ (see Appendix D). Both functions are linear over $(0,1)$. By taking the derivative of $U'_1(P;2)$ and applying the inequalities

$$\frac{\rho_1}{\alpha} > V^\circ > \frac{\gamma_2(\rho_2 - \eta_2) - q_2 d^{\tau_2}\lambda_2}{1 - r_2 d^{\tau_2}} \quad ,$$

it can be shown that

$$\text{for all} \quad P \in (0,1):$$

$$\delta'_1(P;2) > 0 \quad .$$

Similarly,

$$\text{for all} \quad P \in (0,1):$$

$$\delta'_1(P;1) < 0 \quad .$$

Therefore, the required properties are satisfied in F'_1.

We shall now assume that these properties are satisfied in F'_{n-1}, where $n \geqslant 2$; moreover, we shall show that they are also satisfied in F'_n. To do this, we must first consider the special modified game F''_n. In F''_n, the evader is not allowed to move until two looks have been made; that is,

$$F''_n \xrightarrow{\text{look}} F'_{n-1} \xrightarrow{\text{look}} F_{n-2} \quad .$$

Let $U''_n(P;ij)$ represent the payoff in F''_n when the evader uses an optimum strategy and the searcher an optimum strategy after looking first into box i and then into box j. No moving can occur until two looks have been made. Therefore, both $U''_n(P;12)$ and $U''_n(P;21)$ can be expressed as functions of the first two looks plus

$$[Pr_1 + (1 - P)\, r_2]\, d^{\tau_1+\tau_2}\, U_{n-2}\left[\frac{Pr_1}{Pr_1 + (1-P)\, r_2}\right] \quad .$$

Taking the difference between $U''_n(P; 12)$ and $U''_n(P; 21)$ cancels this unknown term, and it can be shown that

for all $P \epsilon (0, 1)$:

$$\delta''_n(P; 21) > \delta''_n(P; 12) \quad .$$

The functions $U''_n(P; 22)$ and $U''_n(P; 21)$ differ only in their dependence on

$$U'_{n-1}\left[\frac{P}{P + (1-P)\, r_2}\, ; 2\right] \text{ and } U'_{n-1}\left[\frac{P}{P + (1-P)\, r_2}\, ; 1\right], \text{ respectively.}$$

For all $P \epsilon (0, 1)$:

$$\delta'_{n-1}(P; 2) > \delta'_{n-1}(P; 1)$$

and it can be shown that

for all $P \epsilon (0, 1)$:

$$\delta''_n(P; 22) > \delta''_n(P; 21) \quad .$$

Using the same reasoning on $U''_n(P; 12)$ and $U''_n(P; 11)$ and combining results, we find that

for all $P \epsilon (0, 1)$:

$$\delta''_n(P; 22) > \delta''_n(P; 21) > \delta''_n(P; 12) > \delta''_n(P; 11) \quad .$$

But,

$$U''_n(P; i) = \min \begin{cases} U''_n(P; i1) \\ \\ U''_n(P; i2) \end{cases}$$

Therefore,

for all $P \epsilon (0, 1)$:

$$\delta''_n(P; 2) > \delta''_n(P; 1) \quad .$$

Consider the functions $U'_{n-1}(P)$ and $U_{n-1}(P)$. In these cases $F''_n \rightarrow F'_{n-1}$ in exactly the same manner as $F'_n \rightarrow F_{n-1}$. Let P_- and P_+ represent the bounding points of the no-move region in F_{n-1}. Furthermore, define $P_-(1)$, $F_-(2)$, $P_+(1)$ and $P_+(2)$ by

$$P_-(1) \xrightarrow{1} P_-$$

$$P_-(2) \xrightarrow{2} P_-$$

$$P_+(1) \xrightarrow{1} P_+$$

$$P_+(2) \xrightarrow{2} P_+ \quad ,$$

and note that $P_-(2) \leqslant P_-(1)$, $P_+(2) \leqslant P_+(1)$.

For all $P \in (0, P_-)$:

$$\delta'_{n-1}(P_-)- \geqslant \delta_{n-1}(P) = \mu_1 \geqslant \delta'_{n-1}(P_-)+ \quad .$$

For all $P \in (P_+, 1)$:

$$\delta'_{n-1}(P_+)- \geqslant \delta_{n-1}(P) = -\mu_2 \geqslant \delta'_{n-1}(P_+)+ \quad .$$

From this, it follows that

for all $P \in [0; P_-(i)]$:

$$\delta''_n[P_-(i); i]- \geqslant \delta'_n(P; i) = \delta'_n[P_-(i); i]- \geqslant \delta''_n[P_-(i); i]+ \quad ;$$

for all $P \in [P_+(i), 1]$:

$$\delta''_n[P_+(i); i]- \geqslant \delta'_n(P; i) = \delta'_n[P_+(i); i]+ \geqslant \delta''_n[P_+(i); i]+ \quad .$$

Furthermore, if $P_- < P_+$ (i.e., $P_- \neq P_+$), we have

for all $P \in [P_-(i), P_+(i)]$:

$$\delta''_n(P; i) = \delta'_n(P; i) \quad .$$

Both $U'_{n-1}(P)$ and $U_{n-1}(P)$ are convex. Therefore, $\delta'_n(P; i)$ and $\delta''_n(P; i)$ are monotonically nonincreasing functions of P. Furthermore, if $P_i < P_j$,

$$\delta'_n(P_i; k)+ \geqslant \delta'_n(P_j; k)- \quad , \quad \text{etc.}$$

Consider the interval $[0, P_-(2)]$; in this case

$$P \in [0; P_-(2)] \Longrightarrow P \in [0; P_-(1)] \quad .$$

Therefore,

for all $P \in [0, P_-(2)]$:

$$\delta'_n(P; 2) \geqslant \delta''_n[P_-(2); 2]+ \geqslant \delta''_n[P_-(1); 2]- > \delta''_n[P_-(1); 1]- \geqslant \delta'_n(P; 1) \quad .$$

(This statement need only be considered when $P_- > 0$, and hence when $P_-(2) < P_-(1)$.)

Now consider the interval $[P_-(2), P_-(1)]$ and note that this interval may intersect $[P_+(2), 1]$.

For all $P \in [P_-(2), P_-(1)]$:

$$\begin{aligned} \delta'_n(P; 2) &\geqslant \delta'_n[P_-(1); 2]- \geqslant \delta''_n[P_-(1); 2]- \\ &> \delta''_n[P_-(1); 1]- \geqslant \delta'_n[P_-(1); 1]- \\ &= \delta'_n(P; 1) \quad . \end{aligned}$$

Therefore,

for all $\quad P \in [0, P_-(1)]$:

$$\delta'_n(P;2) > \delta'_n(P;1) \quad .$$

A similar development shows that

for all $\quad P \in [P_+(2), 1]$:

$$\delta'_n(P;2) > \delta'_n(P;1) \quad .$$

If $P_-(1) < P_+(2)$, we may write:

for all $\quad P \in [P_-(1), P_+(2)]$:

$$\delta'_n(P;2) = \delta''_n(P;2) > \delta''_n(P;1) = \delta'_n(P;1) \quad .$$

Hence,

for all $\quad P \in (0, 1)$:

$$\delta'_n(P;2) > \delta'_n(P;1) \quad .$$

To complete the proof, we must show that

$$U'_n(0;2) < U'_n(0;1) \quad ,$$

$$U'_n(1;2) > U'_n(1;1) \quad .$$

To prove the first inequality, we can apply the inequalities

$$\frac{\gamma_1 \rho_2}{1 - d^{\tau_1}} = \frac{\rho_2}{\alpha} > V^\circ \geqslant \max_P U_{n-1}(P) \geqslant U_{n-1}(0)$$

$$\geqslant U^\infty(0) = \frac{\gamma_2(\rho_2 - \eta_2) - q_2 d^{\tau_2}\lambda_2}{1 - r_2 d^{\tau_2}}$$

to the expression

$$U'_n(0;1) - U'_n(0;2) = \gamma_1\rho_2 - \gamma_2(\rho_2 - \eta_2)$$

$$+ q_2 d^{\tau_2}\lambda_2 + \left(d^{\tau_1} - r_2 d^{\tau_2}\right) U_{n-1}(0) \quad .$$

The second inequality can be proved by using the equivalent approach (or by switching the labels on the boxes).

Convergence:— We can view the truncated games F_n and F'_n as being those games in which the moving costs are set equal to zero after n looks have occurred. As n increases, the evader must wait longer before he can move at no cost. Hence,

for all $\quad P \in (0, 1) \quad , \quad n \geqslant 0$:

$$U_{n+1}(P) \leqslant U_n(P)$$

and

$$\text{for all} \quad P \in (0,1) \quad , \quad n \geqslant 1:$$

$$U'_{n+1}(P) \leqslant U'_n(P) \quad .$$

Since the evader need incur a moving charge only when it is to his advantage,

$$\text{for all} \quad P \in (0,1) \quad , \quad n \geqslant 1:$$

$$U_n(P) \geqslant U'_n(P) \geqslant U^\infty(P) \quad .$$

For any fixed $P \in (0,1)$ neither $U_n(P)$ nor $U'_n(P)$ can increase with n, and both functions are bounded from below by $U^\infty(P)$. Therefore, both functions must converge in the limit and we can define

$$U(P) \equiv \lim_{n\to\infty} U_n(P) \quad ,$$

$$U'(P) \equiv \lim_{n\to\infty} U'_n(P) \quad .$$

The functions $U_n(P)$ and $U'_n(P)$ are bounded, continuous, and convex. Hence, in the limit, U(P) and U'(P) have these properties also. Furthermore, there must exist a unique $P_0 \in (0,1)$ for which

$$P < P_0 \Longrightarrow U'(P;2) < U'(P;1) \quad ,$$

$$P > P_0 \Longrightarrow U'(P;1) > U'(P;2) \quad .$$

Although $0 < P_0^n < 1$ for all finite n, we cannot automatically infer that $0 < P_0 < 1$ in the limit. To prove that this property exists, we must show that

$$U'(0;2) < U'(0;1) \quad ,$$

$$U'(1;2) > U'(1;1) \quad .$$

These inequalities can be proved true by using the previous method, since the inequalities such as

$$\frac{\gamma_1 \rho_2}{1 - d^{\tau_1}} > V^\circ \geqslant U_{n-1}(0) \geqslant \frac{\gamma_2(\rho_2 - \eta_2) - q_2 d^{\tau_2}\lambda_2}{1 - r_2 d^{\tau_2}}$$

hold in the limit.

Therefore, there exists a pair of functions

$$U(P) \equiv \lim_{n\to\infty} U_n(P) \quad ,$$

$$U'(P) \equiv \lim_{n\to\infty} U'_n P \quad ,$$

that satisfy Eqs. (4-3) and (4-4) where

(1) Both are bounded, continuous and convex,

(2) There exists a unique P_0:

$$0 < P_0 < 1 \quad ,$$

$$P < P_0 \Longrightarrow U'(P;2) < U'(P;1) \quad ,$$

$$P > P_0 \Longrightarrow U'(P;2) > U'(P;1) \quad .$$

Uniqueness:– In F_n, the evader can guarantee a payoff of $U_n(P)$ and the searcher can limit the evader to $U_n(P)$, given P. Hence $U_n(P)$ is the value of F_n, given P. Similar considerations apply in F'_n. For all $n \geq 0$, $P \in (0, 1)$, there exists a minimum probability of detection on the next look that is equal to

$$\min \left\{ \begin{matrix} P_0^n q_1 \\ (1 - P_0^n)\, q_2 \end{matrix} \right\} > 0 \quad .$$

This statement applies in the limit as well, since $0 < P_0 < 1$. Hence,

$$\lim_{n\to\infty} \Pr\{F_n \text{ lasts to } F_0\} = 0 \quad .$$

This equation implies that, for any given $P \in (0, 1)$, the evader can guarantee a payoff $\lim_{n\to\infty} U_n(P)$ and the searcher can limit the evader to this amount in F. Similar considerations apply in F'. Hence, $\lim_{n\to\infty} U_n(P)$ is the value of F, given P, and $\lim_{n\to\infty} U'_n(P)$ is the value of F', given P. Furthermore, there is complete feedback (for any P) in Eqs. (4-3) and (4-4). Therefore, $\lim_{n\to\infty} U_n(P)$ and $\lim_{n\to\infty} U'_n(P)$ are the unique bounded solutions to Eqs. (4-3) and (4-4), and they yield the optimum strategies in F and F'.

Theorem 3.

$P_0 \in (P_-, P_+)$, the no-move region. Assume that $P_0 < P_-$, and define $P_-(1)$ by the relation

$$P_-(1) \xrightarrow{1} P_- .$$

The function U(P) is linear over $(0, P_-)$. Hence, $U'(P;1)$ is linear over $[0, P_-(1)]$, where $P_-(1) > P_-$. But,

$$\text{for all} \quad P \in (P_0, 1): \; U'(P) = U'(P;1) \quad .$$

Therefore, $U'(P)$ is linear over $[P_0, P_-(1)]$, which includes P_- as an interior point. This statement contradicts the definition of P_- in Eq. (4-5).

In the same manner, P_0 cannot be greater than P_+. Therefore, $P_0 \in (P_-, P_+)$.

APPENDIX C
A PROPERTY OF $\{W_i'^r(P)\}$

In Sec. 5.5, the general method for calculating the searcher's optimum strategy in H' (his good strategy in G) was developed. The method was extended to cover the generalized reward structure and discounting in Chapters 6 and 7. One as umption was made which we must now prove, i.e., that

$$-\mu_2 \leqslant \frac{dW_i'^r(P)}{dP} \leqslant \mu_1 \quad \text{for each } \sigma_i^r \quad .$$

The payoff associated with each state in the searcher's Markov process is linear in P. Let

$$\delta_i'^r \equiv \frac{dW_i'^r(P)}{dP} \quad ,$$

$$\delta_i^r \equiv \frac{dW_i^r(P)}{dP} \quad ,$$

$$\delta_i^t \equiv \frac{dW_i^t(P)}{dP} = \frac{dW_i'^t(P)}{dP} = \frac{dU_i(P)}{dP} = \frac{dU_i'(P)}{dP} \quad .$$

The above functions are those that apply when the assumed optimum search strategy is used. The derivatives $\delta_i'^r$ and δ_i^r are associated with σ_i^r, and δ_i^t is associated with σ_i^t. In addition, we can associate all of them with an interval π_i belonging to both the no-move and recurrent regions. We can also define

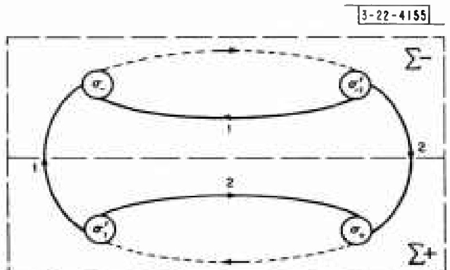
$$\delta_+ \equiv \frac{dW_+(P)}{dP} = \frac{dW_+'(P)}{dP} = \frac{dU_+(P)}{dP} = -\mu_2 \quad ,$$

$$\delta_- \equiv \frac{dW_-(P)}{dP} = \frac{dW_-'(P)}{dP} = \frac{dU_-(P)}{dP} = \mu_1 \quad .$$

Let us first consider the case where both moving regions extend into the recurrent region. Number the intervals (if any) of the no-move region to the left of P_0 as $\pi_{-1}, \pi_{-2}, \ldots$. Number the intervals (if any) that belong to the no-move region and lie to the right of P_0 as $\pi_1, \pi_2, \ldots$. Let π_a (if it exists) represent the interval belonging to the no-move region that is adjacent to π_-. Similarly, let π_b represent the interval in the no-move region that is adjacent to π_+. (Note that π_a and π_b may be the same interval.)

The recurrent chain of the searcher's Markov process can be partitioned into two parts, Σ– and Σ+, and must have the following form:

3-22-4155



The states σ_{-1}^r and σ_- are the same if $P_- = P_0$. States σ_1^r and σ_+ are equivalent if $P_+ = P_0$.

Let us consider the effect that the "good" probability distribution $\underline{Y}_{-1} = \{y_{-1}(1),\ y_{-1}(2)\}$ has on the payoffs associated with each $\sigma_i^r \in \Sigma-$ other than σ_-. If $y_{-1}(2) = 1$, then $\delta_i'^r = \delta_i^t$. If, on the other hand, $y_{-1}(2) = 0$, then $\delta_i'^r = \delta_{i+1}^t$, where $\delta_{i+1}^t \leqslant \delta_i^t$. Since $0 \leqslant y_{-1}(2) \leqslant 1$, we find that $\delta_{i+1}^t \leqslant \delta_i'^r \leqslant \delta_i^t$. But, $-\mu_2 \leqslant \delta_i^t \leqslant \mu_1$, and $-\mu_2 \leqslant \delta_{i+1}^t \leqslant \mu_1$ unless $\sigma_i^r = \sigma_b^r$ when $\sigma_{i+1} = \sigma_+$. In this latter event,

$$\frac{dU_+'(P)}{dP} \leqslant \delta_b'^r \leqslant \delta_b^t, \quad \text{but} \quad \frac{dU_+'(P)}{dP} \leqslant -\mu_2 \quad .$$

Therefore, we must prove that $\delta_b'^r \geqslant -\mu_2$ when $\sigma_b^r \in \Sigma-$.

Similar reasoning can be used to show that if $\sigma_i^r \in \Sigma+$, then $\delta_i^t \leqslant \delta_i'^r \leqslant \delta_{i-1}^t$. Hence, for any state other than σ_a^r belonging to $\Sigma+$, we find that $-\mu_2 \leqslant \delta_i'^r \leqslant \mu_1$, and we must prove that $\delta_a'^r \leqslant \mu_1$ if $\sigma_a^r \in \Sigma+$.

We shall first show that both σ_a^r and σ_b^r must belong to the same set $\Sigma-$ or $\Sigma+$. Assume that $\sigma_b^r \in \Sigma-$. Let $\sigma_i^r \to \sigma_j^r$ imply that a deterministic sequence of looks connects σ_i^r to σ_j^r. Also, let

$$\sigma_i^r \to \sigma_j^r$$

as

$$\sigma_k^r \to \sigma_n^r$$

imply that the same deterministic sequence connects σ_i^r to σ_j^r and σ_k^r to σ_n^r. For the moment, set $y_{-1}(2) = 1,\ y_1(1) = 1$. Since $\sigma_b^r \in \Sigma-$,

$$\sigma_- \to \sigma_b^r \to \sigma_{-1}^r \xrightarrow{2} \sigma_+ \to \sigma_1^r \xrightarrow{1} \sigma_- \quad .$$

But,

$$\sigma_- \to \sigma_b^r \to \sigma_{-1}^r$$

as

$$\sigma_a^r \to \sigma_+ \to \sigma_1^r \quad .$$

Therefore,

$$\sigma_- \to \sigma_a^r \to \sigma_{-1}^r \xrightarrow{2} \sigma_+$$

and $\sigma_a^r \in \Sigma-$. A similar process can be used to show that $\sigma_a^r \in \Sigma- \Longrightarrow \sigma_b^r \in \Sigma-$.

To prove the theorem in the case where both moving regions extend into the recurrent region, we need only show that $\delta_b'^r \geqslant -\mu_2$ if $\sigma_a^r, \sigma_b^r \in \Sigma-$. In general, $\delta_a^t < \delta_- = \mu_1$. If $\delta_a^t = \delta_-$, we are at the lower boundary of a strategy interval, and the proof can be accomplished in the preceding strategy interval. Therefore, we can assume that $\delta_a^t < \delta_- = \mu_1$.

If both σ_a^r and σ_b^r belong to $\Sigma-$, then π_a must lie to the left of P_0 (i.e., $P_- \neq P_0$). If at least two intervals belonging to the no-move region lie to the left of P_0, then $\sigma_{-1}^t \xrightarrow{2} \sigma_+$ and $\sigma_{-2}^r \xrightarrow{2} \sigma_b^r$. Therefore,

$$\sigma_- \to \sigma^r_{-2} \xrightarrow{2} \sigma^r_b$$

as

$$\sigma^t_a \to \sigma^t_{-1} \xrightarrow{2} \sigma_+ \quad .$$

If $\pi_a = (P_-, P_0)$, then $\sigma^r_a = \sigma^r_{-1}$ and

$$\sigma_- \xrightarrow{2} \sigma^r_b$$

as

$$\sigma^t_a \xrightarrow{2} \sigma_+ \quad .$$

In either case,

$$\sigma_- \to \sigma^r_b$$

as

$$\sigma_a \to \sigma_+ \quad .$$

Therefore, $W'_-(P) = W_-(P) = U_-(P)$ is functionally related to $W^r_b(P)$ exactly as $W'^t_a(P) = W^t_a(P) = U_a(P) = U'_a(P)$ is functionally related to $W_+(P) = W'_+(P) = U_+(P)$. We may express these payoffs in the form

$$W_-(P) = a_-P + b_-(1 - P) \quad ,$$

$$W^r_b(P) = a^r_bP + b^r_b(1 - P) \quad ,$$

$$W^t_a(P) = a^t_aP + b^t_a(1 - P) \quad ,$$

$$W_+(P) = a_+P + b_+(1 - P) \quad .$$

The payoff coefficients must be related as follows:

$$a_- = x_a + y_a a^r_b \quad ,$$

$$a^t_a = x_a + y_a a_+ \quad ,$$

$$b_- = x_b + y_b b^r_b \quad ,$$

$$b^t_a = x_b + y_b b_+ \quad ,$$

where y_a is of the form $r^k_1 d^\tau$ and y_b is of the form $r^k_2 d^\tau$ where $k, \tau < \infty$.

If both r_1 and r_2 are greater than zero,

$$\delta^t_a < \delta_- = \mu_1 \Longrightarrow -\mu_2 = \delta_+ < \delta^r_b \quad .$$

If one of the r's is equal to zero, the properties

$$W_a^t(P_-) = W_-(P_-)$$

and

$$W_b^r(P_+) = W_+(P_+)$$

together with

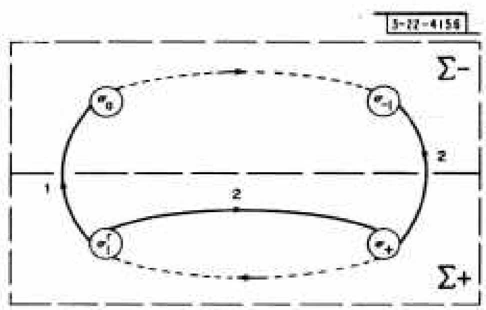
$$\delta_a^t < \delta_- \quad ,$$

$$\Longrightarrow \delta_+ < \delta_b^r \quad .$$

If both r_1 and r_2 are equal to zero, the assumed case in which $P_{01} < P_- < P_0$ cannot occur. Therefore, $\delta_b^r > -\mu_2$. Also, $\delta_b'^r \leqslant \mu_1$. Hence $W_b'^r(P) = W_b^r(P)$, so that $\delta_b'^r = \delta_b^r$. Thus,

$$\delta_b'^r > -\mu_2 \quad .$$

In order to complete the proof, we must consider the case where one but not both of the moving regions extends into the recurrent region. Let us assume that $P_- \leqslant P_{01}$ and $P_+ < P_{02}$. In this case, the recurrent chain must have the following form:



Here, σ_{-1} is a pure state and π_a is the interval lying in the recurrent region with P_{01} as a boundary. For any $\sigma_i \in \Sigma-$, the function $W_i'(P) = U_i'(P)$. Therefore, $\delta_i'^r \neq \delta_i$ only if $\sigma_i^r \in \Sigma+$. For any such state, $\delta_i^t \leqslant \delta_i'^r \leqslant \delta_{i-1}^t$. But σ_a does not belong to $\Sigma+$. Hence, $-\mu_2 \leqslant \delta_i'^r \leqslant \mu_i$ for each recurrent state σ_i^r.

APPENDIX D
SOME PROPERTIES OF G°

We must show that the following conditions exist in G°.

(1) Define $\underline{P}_0$ as a point belonging to the probability simplex that is a solution of the equations

$$U°(\underline{P}; 1) = U°(\underline{P}; 2) = \ldots = U°(\underline{P}; N) \quad .$$

At least one $\underline{P}_0$ must exist and each one that exists must belong to the interior of the simplex.

(2) All boxes are admissible if and only if there exists a $\underline{P}_0$ that is the unique point at which $U°(\underline{P})$ is a maximum. If this occurs, $\underline{P}_0$ must also be the unique solution of

$$U°(\underline{P}; 1) = U°(\underline{P}; 2) = \ldots = U°(\underline{P}; N) \quad .$$

(3) If any inadmissible boxes exist, there must be at least one for which

$$\frac{\rho_i}{\alpha} \leqslant U°(\underline{P}_0) \leqslant \max_{\underline{P}} U°(P) \equiv V° \quad .$$

This statement applies for any $\underline{P}_0$.

(4) In the subsimplex generated by the admissible boxes, there exists a unique $\underline{P}$ where $U°(\underline{P}) = V°$.

(5) There exists a $\underline{Y}$ belonging to the probability simplex with which the searcher can limit the evader to V°. If box i is inadmissible, $y_i = 0$. If box i is admissible, $W°(Y; i) = W°(Y)$.

We have defined $U°(\underline{P})$ by $U(\underline{P}) = \min_i U°(\underline{P}; i)$, where

$$U°(\underline{P}; i) = \frac{\gamma_i \sum_{j=1}^{N} p_j\rho_j - p_i\left(\gamma_i\eta_i + q_i d^{\tau_i}\lambda_i\right)}{1 - d^{\tau_i}(1 - p_i q_i)} \quad .$$

Box i is inadmissible if and only if

$$\frac{\rho_i}{\alpha} \leqslant V° \quad ,$$

and box i is dominated by box j if and only if

$$\frac{\rho_i}{\alpha} \leqslant \frac{\gamma_j(\rho_j - \eta_j) - q_j d^{\tau_j}\lambda_j}{1 - r_j d^{\tau_j}} \quad .$$

Consider some properties of $U°(\underline{P}; i)$. Let S_N represent the probability simplex where for all $i = 1, 2, \ldots, N$, $p_i \geqslant 0$, and $\sum_{i=1}^{N} p_i = 1$. Let $\underline{P}^i$ represent the i^{th} vertex of S_N where $p_i = 1$. A ray belonging to S_N extending from $\underline{P}^i$ intersects the opposite face at a point $\underline{P}'$ where $p_i' = 0$. Along this ray each component of $\underline{P}$ other than p_i can be expressed in the form $p_j = (1 - p_i)\, p_j'$, and

$$\frac{dU^\circ(\underline{P};i)}{dP_i} = \left\{\frac{1}{\left[1-d^{\tau_i}(1-p_iq_i)\right]^2}\right\}$$

$$\left\{-\left(1-r_id^{\tau_i}\right)\gamma_i\sum_{j\neq i}p_j'\rho_j + (1-d^{\tau_i})\left[\gamma_i(\rho_i-\eta_i)-q_id^{\tau_i}\lambda_i\right]\right\} .$$

At any point along this ray except $\underline{P}'$ (and at this point also if $\alpha > 0$), $1-d^{\tau_i}(1-p_iq_i) > 0$. Therefore, along a ray extending from $\underline{P}^i$, $U^\circ(\underline{P};i)$ behaves monotonically. That is, as p_i decreases, $U^\circ(\underline{P};i)$ must be monotonically increasing or decreasing, or equal to a constant. (If box i does not dominate any other box, $U^\circ(\underline{P};i)$ must be monotonically increasing as p_i decreases.)

Let $R_i(c)$ represent the hyperspace $U^\circ(\underline{P};i) = c$. This equation can be expressed in the form

$$p_i\left[\gamma_i(\rho_i-\eta_i)-q_id^{\tau_i}(\lambda_i+c)\right]+\gamma_i\sum_{j\neq i}p_j\rho_j = c \quad .$$

Therefore $R_i(c)$ is a linear hyperplane. If

$$\min_{\underline{P}\epsilon S_N} U^\circ(\underline{P};i) < c < \max_{\underline{P}\epsilon S_N} U^\circ(\underline{P};i) \quad ,$$

then $R_i(c)$ partitions S_N into two nonempty hyperspaces of degree $N-1$. Let us exclude $R_i(c)$ from each of these spaces. Then, $U^\circ(\underline{P};i)$ must be greater than c over one of these spaces and less than c over the other. This follows from the monotonic behavior of $U^\circ(\underline{P};i)$ along any ray in S_N that intersects $\underline{P}^i$. If c is equal to

$$\max_{\underline{P}\epsilon S_N} U^\circ(\underline{P};i) \quad \text{or} \quad \min_{\underline{P}\epsilon S_N} U^\circ(\underline{P};i) \quad ,$$

then $R_i(c)$ must include at least one vertex of S_N. Therefore, $U^\circ(\underline{P};i)$ achieves its maximum over S_N at least one vertex and also its minimum at at least one vertex.

Let $\{A, A'\}$ represent a partition of the boxes into two sets. Let S_A represent the subsimplex of S_N where $\sum_{i\epsilon A} p_i = 1$. Define T_A as the hyperspace belonging to S_N where

$$\text{for all} \quad i, j \in A: \; U^\circ(\underline{P};i) = U^\circ(\underline{P};j) = U^\circ(\underline{P};A) \quad .$$

Let us show that there exists at least one $\underline{P}_0$ belonging to S_N where

$$\text{for all} \quad i = 1, 2, \ldots, N \quad , \quad U^\circ(\underline{P}_0;i) = U^\circ(P_0) \quad .$$

Such a point must belong to the interior of S_N, for at any point P belonging to a boundary of S_N, there must exist at least one $p_i = 0$ and one $p_j > 0$. In this situation, $U^\circ(\underline{P};i) > U^\circ(\underline{P};j)$.

We shall first prove that this property is satisfied when $N = 2$. The simplex S_2 is then the unit interval on the real line, and we may write

$$U^\circ(\underline{P}^1;1) = \frac{\gamma_1(\rho_1-\eta_1)-q_1d^{\tau_1}\lambda_1}{1-r_1d^{\tau_1}} < \frac{\rho_1}{\alpha} = U^\circ(\underline{P}^1;2) \quad ,$$

$$U°(\underline{P}^2; 2) = \frac{\gamma_2(\rho_2 - \eta_2) - q_2 d^{\tau_2}\lambda_2}{1 - r_2 d^{\tau_2}} < \frac{\rho_2}{\alpha} = U°(\underline{P}^2; 1) \quad .$$

As a result, there must exist at least one $\underline{P}_0$ belonging to S_2.

We can use induction to prove that at least one $\underline{P}_0$ exists in S_N. Let A include the first $N - 1$ boxes in a set of N boxes. Assume that there exists a $\underline{P}_{0A}$ belonging to S_A where

$$\text{for all} \quad i \in A: \quad U°(\underline{P}_{0A}; i) = U°(\underline{P}_{0A}; A) \quad .$$

Hence, we assume that T_A intersects S_A. Also,

$$\text{for all} \quad i \in A: \quad U°(\underline{P}^N; i) = \frac{\rho_N}{\alpha} \quad .$$

Therefore, T_A contains the vertex $\underline{P}^N$. Even when $\alpha = 0$, the point $\underline{P}^N$ is a bounding point of T_A, for $U°(\underline{P}; i)$ is bounded and continuous over the interior of S_N.

A simple manipulation reveals that

$$U°(\underline{P}^N; A) > U°(\underline{P}^N; N)$$

and

$$U°(\underline{P}_{0A}; A) < U°(\underline{P}_{0A}; N) \quad .$$

Therefore, there must be at least one point $\underline{P}_0$ belonging to T_A where

$$U°(\underline{P}_0; N) = U°(\underline{P}_0; A) = U°(\underline{P}_0) \quad .$$

Let us assume that all boxes are admissible. Then,

$$\text{for all} \quad i = 1, 2, \ldots, N: \quad \frac{\rho_i}{\alpha} > \max_{\underline{P} \epsilon S_N} U°(\underline{P}) \equiv V° \quad .$$

Consider a point $\underline{P}_0$ belonging to S_N where

$$U°(\underline{P}_0) = V' \leqslant V° \quad .$$

The intersection of $R_i(V')$ with S_N includes the interior point $\underline{P}_0$ and must partition S_N into two nonempty hyperspaces of degree $N - 1$. But,

$$\text{for all} \quad j \neq i: \quad U°(\underline{P}^j; i) = \frac{\rho_j}{\alpha} > V° \geqslant V' \quad .$$

Therefore, all but the i^{th} vertex are included in one hyperspace, and $\underline{P}^i$ is included in the other. Over the latter, $U°(\underline{P}; i) < V'$. This inequality is true for any i and implies that for any $\underline{P} \neq \underline{P}_0$ belonging to S_N, there exist an i and j where $U°(\underline{P}; i) < V' < U°(P; j)$. Therefore, $\underline{P}_0$ must be the unique intersection in S_N of $\{U°(\underline{P}; i)\}$. It must also be the unique point at which $U°(\underline{P})$ is a maximum.

Before proving the converse of this theorem, it is necessary to develop some additional properties concerning $\{U°(\underline{P}; i)\}$. Consider a partition $\{A, A'\}$ where A contains more than one box.

In this case, $\underline{P}^i \in T_A$ for all $i \in A'$. Also, there exists at least one point $\underline{P}_{0A}$ belonging to $T_A \cap S_A$. Consider the hyperplane $R_A(c) \equiv \bigcap_{i \in A} R_i(c)$. If

$$\min_{\underline{P} \in T_A} U^\circ(\underline{P}; A) < c < \max_{\underline{P} \in T_A} U^\circ(\underline{P}; A) \quad ,$$

then $R_A(c)$ partitions T_A into two nonempty hyperspaces. If the boundary $R_A(c)$ is excluded from both of these hyperspaces, $U^\circ(\underline{P}; A)$ is greater than c over one and less than c over the other. It follows that $U^\circ(\underline{P}; A)$ must be a maximum at a vertex $\underline{P}^i \in S_A$ or at a point belonging to $T_A \cap S_A$. Similarly, it must be a minimum at at least one such point.

Let us assume that there exists a $\underline{P}_0$ that is the unique point belonging to S_N at which $U^\circ(\underline{P}) = V^\circ$. Let A include all but the N^{th} box. Let $\underline{P}_{0A}$ represent a point belonging to $T_A \cap S_A$. The function $U^\circ(\underline{P}_{0A}; N) > U(\underline{P}_{0A}; A)$. Therefore, $U^\circ(\underline{P}_{0A}; A) < U^\circ(\underline{P}_0; A) = V^\circ$. This is true for any $\underline{P}_{0A}$ belonging to $T_A \cap S_A$. Hence, $U^\circ(\underline{P}_0; A)$ must be a maximum at $\underline{P}^N$. The hyperplane $R_A(V^\circ)$ cannot contain $\underline{P}^N$ as well as $\underline{P}_0$, for it would then be of degree one and intersect S_A. Therefore,

$$\frac{\rho_N}{\alpha} = U^\circ(\underline{P}^N; A) > U^\circ(\underline{P}_0; A) = V^\circ \quad .$$

Similarly, $\rho_i/\alpha > V^\circ$ for each box, and all boxes must be admissible. Therefore, if $U^\circ(\underline{P})$ is a maximum at a unique point that is some $\underline{P}_0$, all boxes are admissible and $\underline{P}_0$ is the unique point where, for all i, $U^\circ(\underline{P}; i) = U^\circ(\underline{P})$.

Let us consider the case where at least one box is inadmissible. We must first show that $U^\circ(\underline{P})$ will be a maximum at at least one point where, for some i, $U^\circ(\underline{P}; i) > U^\circ(\underline{P})$. Let B represent the set of admissible boxes and B' the set of inadmissible boxes. Then,

$$\text{for all} \quad i \in B: \quad \frac{\rho_i}{\alpha} > V \quad ;$$

$$\text{for all} \quad i \in B': \quad \frac{\rho_i}{\alpha} \leqslant V \quad .$$

Consider a point $P \in S_B$.

$$\text{For all} \quad j \in B': \quad U^\circ(\underline{P}; j) > \min_{i \in B} U^\circ(\underline{P}; i) \quad .$$

Let $\underline{P}_{0B}$ be a point belonging to S_B at which

$$\min_{i \in B} U^\circ(\underline{P}_{0B}; i) = \max_{\underline{P} \in S_B} \{\min_{i \in B} U^\circ(\underline{P}; i)\} \equiv V_B^\circ \quad .$$

Then, $V_B^\circ \leqslant V^\circ$ and

$$\text{for all} \quad i \in B: \quad \frac{\rho_i}{\alpha} > V^\circ \geqslant V_B^\circ \quad .$$

Therefore, all boxes belonging to B are also admissible in the reduced game involving only those boxes. Hence, $\underline{P}_{0B}$ is the unique point belonging to $T_B \cap S_B$.

Consider a $\underline{P}_0 \in S_N$ where for all $i = 1, 2, \ldots, N$, $U^\circ(\underline{P}_0; i) = U^\circ(\underline{P}_0)$, and assume that $V_B^\circ < U^\circ(P_0)$. Define B" as the subset of B' where

$$\text{for all } i \in B'': \quad \rho_i = \max_{j \in B'} \rho_j$$

Then, $U^\circ(\underline{P}; B)$ is a maximum over $S_{B''}$ which does not include $\underline{P}_0$. Therefore, for any $i \in B''$,

$$U^\circ(\underline{P}_0) = U^\circ(\underline{P}_0; B) < \frac{\rho_i}{\alpha} \leqslant V^\circ \quad .$$

If V_B° is less than $U^\circ(\underline{P}_0)$, then $U^\circ(\underline{P})$ can not be a maximum at $\underline{P}_0$.

If $U^\circ(\underline{P}_0) = V_B^\circ$, then $U^\circ(\underline{P}_0) \leqslant V^\circ$. In this situation, $U^\circ(P)$ will be a maximum either at $\underline{P}_{0B}$ as well as $\underline{P}_0$ or at neither point. Therefore, if any inadmissible boxes exist, $U^\circ(\underline{P})$ must be a maximum at at least one point where all the functions $\{U^\circ(\underline{P}; i\}$ do not intersect.

Assume that $U^\circ(\underline{P}) = V^\circ$ at a point where

$$\text{for all } j \neq i: \quad U^\circ(\underline{P}; j) > U^\circ(\underline{P}; i) = U^\circ(\underline{P}) \quad .$$

But, $U^\circ(\underline{P}; i)$ is a maximum at at least one vertex, and $\underline{P}^i$ is the only vertex where $U^\circ(\underline{P}; i) = U^\circ(\underline{P})$. Therefore, $U^\circ(\underline{P})$ is a maximum at $\underline{P}^i$ and box i dominates all of the other boxes. The function $U^\circ(\underline{P}; i)$ must be a minimum at at least one vertex other than $\underline{P}^i$, and there exists at least one inadmissible box where

$$\frac{\rho_j}{\alpha} \leqslant U^\circ(\underline{P}_0) \leqslant V^\circ \quad .$$

If no box dominates all the others, and if at least one box is inadmissible, $U^\circ(\underline{P})$ must be a maximum at at least one point belonging to a subspace T_A. Here, A must include at least two boxes, but we can choose it so that it does not include all N. Furthermore, the point in T_A can be chosen so that

$$\text{for all } i \notin A: \quad U^\circ(\underline{P}; i) > U^\circ(\underline{P}; A) = V^\circ \quad .$$

The function $U^\circ(\underline{P}; A)$ must be a maximum at a point belonging to $T_A \cap S_A$ or at a vertex $\underline{P}^i$ not belonging to S_A. At each of these vertexes, $U^\circ(\underline{P}^i; A) > U^\circ(\underline{P}^i; i)$. Therefore, there exists a point $\underline{P}_{0A} \in T_A \cap S_A$ where $U^\circ(\underline{P}_{0A}; A) = V^\circ$. Also,

$$\text{for all } i \notin A: \quad \frac{\rho_i}{\alpha} = U^\circ(\underline{P}^i; A) \leqslant U^\circ(\underline{P}_{0A}; A) = V^\circ \quad .$$

Hence, A must contain all of the admissible boxes, and for any P_0 in S_N there must be at least one $i \in A$ where

$$\frac{\rho_i}{\alpha} \leqslant U^\circ(\underline{P}_0) \leqslant V^\circ \quad .$$

The set A could have been chosen so that it would contain some inadmissible boxes. However, it contains all of the admissible boxes and there exists a $\underline{P}_{0A} \in T_A \cap S_A$ where $U^\circ(\underline{P}_{0A}) = V^\circ$. Therefore, we can consider the reduced game involving only those boxes belonging to A, and repeat the process. Each time this is done, at least one inadmissible box is removed. Eventually only B, the set of admissible boxes, can remain. In S_N, the intersection $T_B \cap S_B$ consists of a unique point $\underline{P}_{0B}$, and we can now state that $U^\circ(\underline{P}_{0B}; B) = U^\circ(\underline{P}_{0B}) = V^\circ$. Also, for any $\underline{P}_0$ in S_N, there must exist at least one inadmissible box for which

$$\frac{\rho_i}{\alpha} \leqslant U^\circ(\underline{P}_0) \leqslant V^\circ \quad .$$

Let us now show that there exists a good strategy for the searcher. This will imply that $V°$ is the value of $G°$. Assume that the evader first hides with probability $\underline{P} \in S_N$ and that the searcher looks first into a box i that belongs to B. Assume further that the evader always uses the optimum $\underline{P}_{0B}$ from then on and that the searcher never looks into an inadmissible box. Let $U°'(\underline{P}; i)$ represent the resulting payoff:

$$U°'(\underline{P}; i) = \gamma_i \sum_{j=1}^{N} p_j\rho_j - p_i\left(\gamma_i\rho_i + q_i d^{\tau_i}\lambda_i\right) + (1 - p_i q_i) d^{\tau_i} V° \quad .$$

But, $U°'(\underline{P}; i)$ is linear in $\underline{P}$ over S_B. Furthermore,

$$\text{for all} \quad i \in B: \quad U°'(\underline{P}_{0B}; i) = V° \quad .$$

Since $\underline{P}_{0B}$ is the unique point in S_B where $\min_{i \in B} U°(\underline{P}; i) = V°$,

$$\text{for all} \quad i \in B: \quad U°(\underline{P}^i; i) < V° \quad .$$

As a result, there must exist a unique probability vector $\underline{Y}_0 \in S_B$ where

$$\text{for all} \quad \underline{P} \in S_B: \quad \sum_{i \in B} y_{0i} U°'(\underline{P}; i) = V° \quad .$$

If the searcher uses $\underline{Y}_0$ to determine each look, the payoff will equal $V°$ as long as the evader never hides in an inadmissible box. If he does hide in such a box, the payoff cannot be greater for $\rho_i/\alpha \leqslant V°$ for each such box. Therefore, $\underline{Y}_0$ limits the evader to $V°$, a payoff that the evader can guarantee. The searcher's good strategy is defined by $\underline{Y}_0$, which can be calculated by using the techniques suggested in Sec. 8.3.

UNCLASSIFIED

UNCLASSIFIED