Journal of the Korean
OR/ MS Society
Vol. 19, No. 1, April 1994

A MARKOV DECISION PROCESSES FORMULATION FOR THE LINEAR SEARCH PROBLEM

Z.T. Balkhi* and L. Benkherouf*

Abstract

The linear search problem is concerned with finding a hiden target on the real line **R**. The position of the target is governed by some probability distribution. It is desired to find the target in the least expected search time. This problem has been formulated as an optimization problem by a number of authors without making use of Markov Decision Processes (MDP) theory. It is the aim of the paper to give a (MDP) formulation to the search problem which we feel is both natural and easy to follow.

1. Introduction

For simplicity, we shall only consider the simplest linear search model where the starting search point is taken to be the origin. Assume that a target P is hidden on the real line **R** such that its position is given by a random variable X whose distribution is denoted by F(x) as depicted in Figure 1.

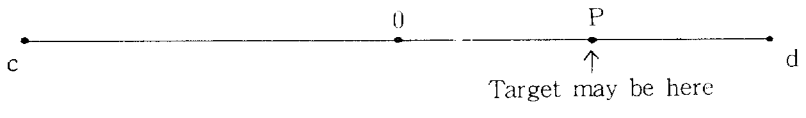

Figure 1. Showing the positon of the target

A searcher starts looking for the target from some starting point, 0 say, and moves continuously in both direction until the target is found. The objective of the searcher is to find the target so as to minimize the expected length of the search path. This objective is assumed to

* Department of Statistics and Operations Research, King Saud University, Riyadh 11451, Saudi Arabia

be equivalent to finding the target in the least expected time. To be more precise, let

$$c = \inf\{t: F(t)>0\},\ d = \inf\{t: F(t)<1\}. \qquad (1)$$

The searcher starts from position 0, he either chooses to move to the right (or the left) of 0 a distance a_1. If the target is not found, he moves in the opposite direction a distance a_2 from 0 and if the target is not found he again moves in the opposite direction a distance a_3 where $a_3 > a_1$. This process is repeated until the target is found. So a policy can be represented by a sequence $\{a_i, i \geq 0\}$ such that

$$\cdots a_5 \leq a_3 \leq a_1 \leq a_0 = 0 \leq a_2 \leq a_4 \leq a_6 \leq \cdots \qquad (2)$$

with $a_{2i-1} \to c$ and $a_{2i} \to d$ or

$$\cdots a_6 \leq a_4 \leq a_2 \leq a_0 = 0 \leq a_1 \leq a_3 \leq a_5 \leqslant \cdots \qquad (3)$$

with $a_{2i-1} \to d$ and $a_{2i} \to c$.

Expression (2) corresponds to the first movement of the searcher is to the left of 0 while expression (3) means that the first movement is to the right of 0.

Recall that x denotes the position of the target and define D(a,x) to be the total distance travelled from the starting point 0 to x using a search path $a = \{a_i, i \geq 0\}$. Then, by letting $a_{-1} = 0$, we have

$$D(a,x) = |x| + 2\sum_{i=1}^{n-1} |a_i|$$

where $a_{n-2} \leq x \leq a_n$. It follows that the expected travelled distance to find the target is then

$$D(a,F) = \int_c^d |x|\, dF(x) + 2\sum_{n=1}^{\infty}\sum_{i=1}^{n-1} |a_i| \int_{a_{n-2}}^{a_n} dF(x),$$

which can be reduced to

$$D(a,F) = \int_c^d |x|\, dF(x) + 2\sum_{i=1}^{\infty} |a_i| \left(1 - \int_{a_{i-1}}^{a_i} dF(x)\right) \qquad (4)$$

The above derivation is due to Beck [3] and Franck [4]. The problem of finding the optimal search path then reduces to the problem of minimizing expression (4). For treatment of the problem along these lines please consult(Balkhi [1],[2]). However, a close look at the way the formulation was derived reveals that the learning about the position of the target has not been incorporated in the prior probability representing the position of the target. For example, if a search of distance a_1 from 0 has been carried out, our knowledge about the distribution F(x) should change but this seems to have been ignored here. In the next section we shall give a MDP formulation to the problem which is simple and takes the learning process into account and which we feel is both natural and easy to follow. Also, we hope that this will be a strating point for trying to find optimal search strategies based on the MDP formulation.

2. The Markov Decision Processes Formulation

Here, we shall represent the state of the system at any time by the vector (x,y). This mean that x distance to the right of 0 was searched and y distance to the left of 0 was searched and no target was found. Let $V^L(x,y)$ denotes the minimum expected search distance to find the target given that the searcher is to the left of 0 (at position y) and the state of the system is (x,y). The expression $V^R(x,y)$ has the same meaning except that the searcher is now to the right of 0 (at position x). If we let V(0,0) denote the minimum expected search distance to find the target, then the optimality equation can be written as

$$V(0,0) = \min \{V^L(0,0),\ V^R(0,0)\} \tag{5}$$

and

$$V^R(x,y) = \min_{T>y} \{x+T+(1-\int_{-T}^{x} dF(p))(1-\int_{-y}^{x} dF(p))^{-1}\ V^L(x,T)\}, \tag{6}$$

$$V^L(x,y) = \min_{Z>x} \{Z+y+(1-\int_{-y}^{Z} dF(p))(1-\int_{-y}^{x} dF(p))^{-1}\ V^R(Z,y)\}, \tag{7}$$

where,

$$(1-\int_{-T}^{x} dF(p))(1-\int_{-y}^{x} dF(p))^{-1} \tag{8}$$

is the conditional probability that the target is not between $-T$ and x given that it is not between $-y$ and x. The term

$$(1-\int_{-y}^{Z} dF(p))(1-\int_{-y}^{x} dF(p))^{-1}$$

has a similar interpretation.

Note that relation (6) means that we are currently positioned to the right of 0 given that the state of the systm is represented by (x,y) and the actions available are to move some distance T (to be determined) to the left of zero. If the target is found then we stop seaching. Otherwise we continue searching from our new position (x,T) situated to the left of zero noting that the probability of not finding the target is given by (8). Relation (7) has a similar interpretation as that of (6). Relation (5) comes from the fact that initially we have the choice of moving either to the left of zero or to the right of zero. Also, note that the set of actions available to the searcher at any postion (x,y) consists of the open intervals (x,d) and (c,y), which are continuous, where c and d are given by (1).

The formulation based on relations (5), (6) and (7) is a (MDP) formulation since the choice of our actions at any time is dictated only by our current state (position) (x,y) and does not depend on the past. This means that all the theory related to (MDP) can be applied to the search problem. Of particular interest are the computational techniques used for solving (MDP). This will make the problem easier to handle computationally than previously was. For more details about the techniques consult Ross [5]. In the example to be presented later, we used the method of successive approximation with a starting solution $V^R(x,y) = V^L(x,y) = 0$ for all x and y.

Note that it can be shown that the problem of finding the optimal search path with the (MDP) formulation can be stated equivalently as the problem of finding the sequence $a = \{a_i, i \geq 0\}$ that minimizes the following expression

$$R(a,F) = \sum_{i=1}^{\infty} |a_i| \int_{a_{i-2}}^{a_i} dF + 2 \sum_{i=1}^{\infty} |a_i| \left(1 - \int_{a_{i-1}}^{a_i} dF\right)$$

with $a_{-1} = a_0 = 0$. To see that, assume without loss of generality that the first movement a_1 occurs to the right of 0, then we can write

$$R(a,F) = a_1 + \left(1 - \int_0^{a_1} dF(p)\right) V^R(a_1, 0)$$

The second term on the right hand side of the above expression comes from the fact we searched a distance a_1 to the right and the target was not found. This occurs with probability $1 - \int_0^{a_1} dF(p)$. Then, we continue searching from our new position situated which is situated a distance a_1 to the right of 0. Therefore, if we move a_2 to left of 0 we get:

$$R(a,F) = a_1 + \left(1-\int_0^{a_1} dF(p)\right) \left\{ |a_1| + |a_2| \left(1-\int_{a_2}^{a_1} dF(p)\right) \left(1-\int_0^{a_1} dF(p)\right)^{-1} V^L(a_1, a_2) \right\},$$

repeating this process indefinitely gives:

$$R(a,F) = a_1 + \sum_{i=1}^{\infty} \left(1-\int_{a_{i-1}}^{a_i} dF(p)\right) \left(|a_i| + |a_{i+1}|\right)$$

which leads after some algebraic manipulation to the required result.

A comparison between relations (4) and (5) reveals that their diference is

$$R(a,F)-D(a,F) = \sum_{i=1}^{\infty} |a_i| \int_{a_{i-2}}^{a_i} dF - \int_c^d |x| \, dF(x) \tag{9}$$

This difference is non negative since

$$\int_c^d |x| \, dF = \sum_{i=1}^{\infty} \int_{a_{i-2}}^{a_i} |x| \, dF \leq \sum_{i=1}^{\infty} |a_i| \int_{a_{i-2}}^{a_i} dF.$$

In fact relation (9) indicates that the MDP formulation and the classical formulation may lead to different search strategies. It also points to the fact that the difference between their expected distance travelled is bounded by:

$$\min_a \{R(a,F)-D(a,F)\},$$

where the above difference is given by (9).

3. Numerical example and conclusion

Example : Assume that the probability describing the position of the particle along the line is standard normal. Note that because of the symmetry of the distribution at 0, it is not important in which direction to start. Accordingly, we assume that the first move occurs to the right of 0. The classical method based on relation (4) gives the following search sequence:

$a_1=1.4,\ a_2=2.6,\ a_3=1.4,\ a_3=3.6,\ a_4=4.9,\ a_5=5.3,\ a_6=6.0$

with a total expected distance travelled of 2.99.

On the other hand, the MDP formulation gives:

$$a_1=1,\ a_2=1.7,\ a_3=2.2,\ a_3=2.9,\ a_4=3.4,\ a_5=6.0,\ a_6=6.0$$

with a total expected distance travelled $V(0,0)=3.99$

To sum up, in this note we formulated the classical linear search problem as a MDP. By doing this we are hoping to generate interest in devising optimal search strategies based on this formulation. We also presented a numerical example to show that the classical formulation and the MDP formulation may lead to different policies.

References

[1] Balkhi, Z.T., "The Generalized Linear Search Problem, Existence of Optimal Paths," *Journal of Operational Research Society of Japan*, Vol.30(1987), pp.399-421.

[2] Balkhi, Z.T., "Generalized Optimal Search Paths for Continuous Univariate Random Variables," *R.A.I.R.O.-Rech.O.*, Vol.23, No.1(1989), pp.67-96.

[3] Beck, A., "On the Linear Search Problem," *ISR. J. Math.*, Vol.2(1964), pp.221-228.

[4] Frank, W., "An Optimal Search Problem," *SIAM Rev.*, Vol.7(1965), pp.503-512.

[5] Ross, S., *Applied Probability Models with Optimization Application*, Holden Day, San Francisco, 1970.