Optimal Search for a Moving Target with the Option to Wait

János Flesch,[1] Emin Karagözoğlu,[2] Andrés Perea[1]

[1] *Department of Quantitative Economics, Maastricht University, 6200 MD Maastricht, The Netherlands*

[2] *Department of Economics, Maastricht University, 6200 MD Maastricht, The Netherlands*

Abstract: We investigate the problem in which an agent has to find an object that moves between two locations according to a discrete Markov process (Pollock, *Operat Res* 18 (1970) 883–903). At every period, the agent has three options: searching left, searching right, and waiting. We assume that waiting is costless whereas searching is costly. Moreover, when the agent searches the location that contains the object, he finds it with probability 1 (i.e. there is no overlooking). Waiting can be useful because it could induce a more favorable probability distribution over the two locations next period. We find an essentially unique (nearly) optimal strategy, and prove that it is characterized by two thresholds (as conjectured by Weber, *J Appl Probab* 23 (1986) 708–717). We show, moreover, that it can never be optimal to search the location with the lower probability of containing the object. The latter result is far from obvious and is in clear contrast with the example in Ross (1983) for the model without waiting. © 2009 Wiley Periodicals, Inc. Naval Research Logistics 56: 526–539, 2009

Keywords: search for a moving target; waiting; Markovian dynamics

JEL Classification Codes: C44, C61, C72

1. INTRODUCTION

1.1. Description of the Problem without Waiting

In 1970, Pollock considered the following search problem: there is an object moving across two locations labeled as L (left) and R (right) according to a discrete time Markov process (see Fig. 1).

Hence, if the object currently is at L, it will move to R with probability q. If the object currently is at R, it will move to L with probability r. The probabilities q and r are called the transition probabilities. The probability that the object is initially at L is $\bar{p}$, and hence $1 - \bar{p}$ is the probability that the object initially is at R.

The periods at which the agent can search for the object are denoted by $t = 1, 2, 3, \ldots$ As soon as the agent finds the object, the process stops. If the object has not yet been found, the agent has two possible actions at any given period: searching left (L) and searching right (R), which have costs c_L and c_R, respectively. Assume that the object is overlooked with probability α_L if it is located at L and the agent searches L, and overlooked with probability α_R if it is located at R

Correspondence to: E. Karagözoğlu (E.Karagozoglu@ALGEC. unimaas.nl)

and the agent searches R. The objective is to minimize the total expected searching cost needed for finding the object.

1.2. Overview of the Literature

Ross [7] conjectured that there exists an optimal searching strategy of the following form: search L at period t if and only if $p_t \geq \tilde{\pi}$, where p_t is the probability that the object is located at L at period t. Here, $\tilde{\pi}$ is a threshold, which depends on the parameters of the model (i.e., q, r, and $\bar{p}$). The conjecture is intuitive, but Ross proved it only for the special case, $c_L = c_R$. It is intriguing that despite extensive work by numerous researchers, this conjecture, in its full generality, remains unproved to this day.

Pollock [6] showed the existence of such a threshold $\tilde{\pi}$ for the special case without overlooking (i.e., $\alpha_L = \alpha_R = 0$) in the discrete time model. He also computed the threshold $\tilde{\pi}$ for every possible configuration of parameters. Weber [10] proved Ross' conjecture in full generality, but in a continuous time (rather than discrete time) setting. He also conjectured that for the case with waiting, the optimal strategy can be characterized by two thresholds. White [11] proved Ross' conjecture for a fairly wide range of circumstances and provided a mathematical structure on which, further work on the
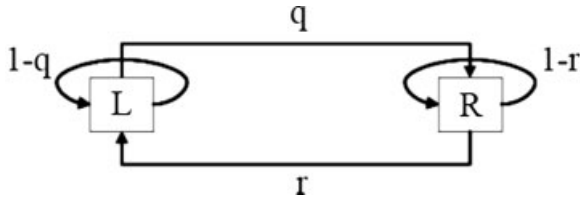
Figure 1. Search for a moving target.

problem can be based. Later, MacPhee and Jordan [4] proved Ross' conjecture for a discrete time Markov process with the possibility of overlooking, but only for some parameter configurations.

Assaf and Sharlin-Bilitzky [1] investigated a continuous time version of the model in which the agent, in addition, can specify the searching effort. The main result of their article is to describe the optimal searching strategy in this model. Kan [3] investigated the search problem above with the additional option to stop searching. The article characterized optimal strategies for specific classes of transition probabilities.

1.3. Our Contribution

Our article extends the original search problem by introducing the option to wait.[1] That is, at every period the agent has now three options: searching left (L), searching right (R), and waiting (W). We assume that waiting is costless whereas searching is costly. This extension is motivated by the fact that an agent can benefit from waiting by getting a more favorable probability distribution over the two locations next period.

We assume that the costs of searching left and right are equal (normalized to 1) and the overlooking probabilities are 0. If the agent finds the object he receives a prize P. Provided the prize is high enough, this gives the agent an incentive to find the object with probability 1. Without this prize, it would be optimal to simply wait forever. In line with the previous literature mentioned above, we assume that future payoffs are not discounted.

We prove that, if $P \geq 2$ and the sum of the transition probabilities (i.e., $q + r$) is at least 1, or $q = r$, then there exists an optimal strategy of the following form: search R if $p_t \in [0, \pi_1]$, wait if $p_t \in (\pi_1, \pi_2)$ and search L if $p_t \in [\pi_2, 1]$. (Recall that p_t denotes the probability that the object is at location L at period t, given it has not been found yet.) That is, this optimal strategy is given by two thresholds, π_1 and π_2, as conjectured by Weber [10]. Moreover, this optimal strategy is unique up to the choices at π_1 and π_2.

On the other hand, if $P \geq 2$ and $0 < q + r < 1$ and $q \neq r$, we prove that for every $\varepsilon > 0$ there exists an ε-optimal

strategy of the form described above. By an ε-optimal strategy we mean that the agent cannot improve his expected utility by more than ε by switching to another strategy. In this case, an optimal strategy need not exist since there are situations in which waiting for one more period always provides more favorable probabilities p_t and $1 - p_t$ for the next period. In these situations, the agent wants to postpone searching as long as possible.

We show that our thresholds π_1 and π_2 have an interesting relationship with the threshold $\tilde{\pi}$ as derived by Pollock for the model without waiting: It turns out that $\tilde{\pi}$ is always in between π_1 and π_2. Consequently, if the agent searches a location in the model with waiting, he would search the same location in the model without waiting. Although this result is not unexpected, it is not immediately clear why it should be the case.

In the model without the option to wait, Ross [7] has shown that it is possible that searching the location with the lower probability of containing the object is optimal. The reason is that searching this location may induce a very favorable probability distribution for the location of the object next time, in case it is not found. We show that this is no longer possible in our model: the agent either searches the location with the higher probability of containing the object, or he waits. So, even if the lower probability is close to $\frac{1}{2}$, and searching the location with the lower probability would provide the agent with complete certainty next period, he would still not search that location. This fact is far from obvious.

We would like to make some remarks on the model. Although the optimal strategies above are designed for an infinite horizon model without discounting, they remain "nearly optimal" if we either (1) apply them to the discounted case with a discount factor close enough to 1, or (2) use them in the finite horizon case with sufficiently many periods. The reason is that after sufficiently many periods the strategies we propose will have found the object with probability almost 1. In this respect the optimal strategies we construct are robust, and particularly useful in situations where either the discount factor is large but not completely known, or there is a large but unknown horizon.

From a mathematical point of view, the undiscounted case raises additional difficulties because the expected payoff is not continuous in the strategy being chosen. For instance, if the object does not move (that is, $q = r = 0$) and is at L with probability 1, any strategy that waits for finitely many periods and then searches L is optimal. However, the "limit strategy" would wait for an infinitely long time, and is therefore not optimal. As a consequence of this type of discontinuity, optimal strategies may not even exist in general, as we already discussed above.

Finally, for a game-theoretic discussion of a similar search problem with multiple searchers, we refer the reader to Flesch et al. [2].

---

[1] Note that waiting is different from stopping analyzed in Kan [3]. The agent can continue searching after waiting, whereas stopping means no more search can be done.

Table 1. Values for the thresholds π_1 and π_2 when $q \leq r$. If $q < r$ then α is the smallest positive probability amongst $\bar{p}, 1 - \bar{p}, q, 1 - q, r,$ and $1 - r$. If $q = r$, then $\alpha = 0$.

	π_1	π_2
$q + r = 0$	$\dfrac{1}{2}$	$\dfrac{1}{2}$
$0 < q + r < 1$ and	$\dfrac{q(1+r)}{(q+r)(1+q)}$	$\max\left\{\dfrac{1}{2}, \dfrac{r}{q+r} - \dfrac{\alpha\varepsilon}{2}\right\}$
$(q+1)r^2 + q^2 r - q \leq 0$ (Area A in Figure 2)	(unique p with $V(p, R) = V(\pi^*, L)$)	
$0 < q + r < 1$ and	$\dfrac{q}{(q+r)(1+q) - q}$	$\max\left\{\dfrac{1}{2}, \dfrac{r}{q+r} - \dfrac{\alpha\varepsilon}{2}\right\}$
$(q+1)r^2 + q^2 r - q > 0$ (Area B in Figure 2)	(unique p with $V(p, R) = V(\pi^*, L)$)	
$q + r = 1$	q	r
$1 < q + r < 2$ and	$\dfrac{q(1-r)}{(q+r)(1-q)}$	$\dfrac{r}{q+r}$
$(1 - q + q^2)r - 2q^2 + q^3 \leq 0$ (Area C in Figure 2)	(unique p with $V(p, R) = V(Ap, L)$)	
$1 < q + r < 2$ and	$\dfrac{1-r}{(q+r)(1-q) + 1 - r}$	$\dfrac{r}{q+r}$
$(1 - q + q^2)r - 2q^2 + q^3 > 0$ (Area D in Figure 2)	(unique p with $V(p, R) = V(Ap, L)$)	
$q + r = 2$	$\dfrac{1}{2}$	$\dfrac{1}{2}$

For our main theorem, any $\tilde{\alpha}$ such that $0 < \tilde{\alpha} < \alpha$ would also do in case $q < r$. The parts between brackets will become clear after reading Section 4.2.

1.4. Road Map

Section 2 of this article introduces the necessary notation and presents our main theorem along with the values of the thresholds π_1 and π_2. Section 3 discusses all different types of Markov processes. Section 4 introduces the value function and analyzes its properties, which we use in the proof of the main theorem. Section 5 contains preparatory lemmas for the proof of the main theorem. Section 6 contains two lemmas which prove the existence of the thresholds for all possible types of Markov processes. In Section 7, we compute the values of the thresholds π_1 and π_2 (see Table 1). In Section 8, we compare our optimal strategy with the optimal strategy for the model without waiting, as derived by Pollock [6]. Section 9 provides a discussion of the assumptions of the model, as well as a comparison of the results from different models.

2. MAIN THEOREM

2.1. Optimal and ε-Optimal Strategies

A strategy γ is defined as a sequence $(\gamma_t)_{t \in \mathbb{N}}$, where $\gamma_t \in \{L, R, W\}$ describes the action to be taken at period t if the object has not been found yet.[2] The expected utility induced by a strategy γ is denoted by $U(\gamma)$, while

[2] We do not consider mixed or history dependent strategies here, since the agent will not be able to improve upon his cost by turning to such strategies.

the expected searching cost is called $C(\gamma)$. Thus, $U(\gamma) = \mu(\gamma)P - C(\gamma)$, where $\mu(\gamma)$ is the probability of finding the object.

We now mention some specific classes of strategies. We call a strategy γ successful if $\mu(\gamma) = 1$. In that case, $U(\gamma) = P - C(\gamma)$. For every period t, let $\mu_t(\gamma)$ be the probability of finding the object at period t, given that it has not been found yet. Also, let $C_t(\gamma)$ be the cost at period t (so, $C_t(\gamma) = 1$ if the agent searches, and 0 if he waits). We call the strategy γ loss-free if $\mu_t(\gamma)P - C_t(\gamma) \geq 0$ for every period t, meaning that at every period the expected utility is non-negative. Moreover, we call γ strictly loss-free if $\mu_t(\gamma)P - C_t(\gamma) > 0$ for every period t at which γ prescribes to search.

Formally, the agent's objective is to find, for a given configuration $(\bar{p}, q, r)$ of initial and transition probabilities, a strategy γ that maximizes the expected utility, $U(\gamma)$. A strategy γ is called optimal if $U(\gamma) \geq U(\gamma')$ for all other strategies γ'. For any $\varepsilon > 0$, the strategy γ is called ε-optimal if $U(\gamma) \geq U(\gamma') - \varepsilon$ for all other strategies γ'. As we will show below, optimal strategies do not always exist for every configuration of initial and transition probabilities, but for every $\varepsilon > 0$ we can always construct an ε-optimal strategy.

We will show that, if $P \geq 2$, then it is sufficient to consider successful strategies.

LEMMA 2.1: Let $P \geq 2$, and let γ be an arbitrary strategy. Then, there is a successful strategy γ' with $U(\gamma') \geq U(\gamma)$.

PROOF: Suppose that γ is not successful. Let $\mathbb{N}(\gamma)$ denote the set of periods at which γ searches a location with probability zero of containing the object. Let $\tilde{\gamma}$ denote the strategy which follows the prescriptions of γ, except that $\tilde{\gamma}$ waits at all periods in $\mathbb{N}(\gamma)$. It is clear that $C(\tilde{\gamma}) \leq C(\gamma)$, and that $\tilde{\gamma}$ finds the object with the same probability as γ. Thus, $U(\tilde{\gamma}) \geq U(\gamma)$, and $\tilde{\gamma}$ is also unsuccessful. Let z denote the smallest positive probability amongst the initial probabilities $\bar{p}, 1 - \bar{p}$ and the transition probabilities $q, 1 - q, r$ and $1 - r$. Note that whenever $\tilde{\gamma}$ searches a location, then this location contains the object with probability at least z. Since $\tilde{\gamma}$ is unsuccessful, this implies that $\tilde{\gamma}$ can only search finitely many times. Thus, there exists a period T such that $\tilde{\gamma}$ only waits at all periods $t \geq T$.

Now, let γ' be the strategy that coincides with $\tilde{\gamma}$ during the first $t - 1$ periods, and which from period t on always searches the location with the highest current probability of containing the object. (If both locations are equally likely, check L). Then, γ' will find the object with probability 1, so is a successful strategy. Suppose that the object has not been found before period t. Then it can be shown by induction on k that the expected searching cost until period $t + k$ will be at most $1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{k-1}} \leq 2$. Therefore, the expected continuation searching cost for γ' from period t on is at most 2. So, the expected continuation utility for γ' from period t on is at least 0, as $P \geq 2$. But then, $U(\gamma') \geq U(\tilde{\gamma}) \geq U(\gamma)$. □

It thus follows that, if $P \geq 2$, we may restrict ourselves to successful strategies. This is useful since finding an optimal strategy within the class of successful strategies is equivalent to minimizing the expected searching costs.

The prize, P, should be at least 2 to make the agent search. If $P < 2$, the agent may find it optimal to wait forever since the benefit from finding the object does not cover the expected costs of search in some cases. Consider, for instance, the situation where the initial probability $\bar{p}$ is $\frac{1}{2}$, and the transition probabilities q and r are $\frac{1}{2}$ as well. In that case, the expected cost of finding the object would be exactly 2.[3] Hence, the agent would have no incentive to search at all.

2.2. Our Main Theorem

Our main theorem states that, whatever configuration of initial probabilities and transition probabilities we take,

we can always construct a (nearly) optimal strategy of the following kind:

- if the probability that the object is at L is below a threshold π_1, search R;
- if the probability that the object is at L is between π_1 and another threshold $\pi_2 \geq \pi_1$, wait;
- if the probability that the object is at L is above π_2, search L.

The (nearly) optimal strategy is therefore completely characterized by the thresholds π_1 and π_2. It is possible that the two thresholds coincide. In that case, the agent will never wait.

In the main theorem, let $p_t(\gamma)$ be the probability that at period t the object is at L if strategy γ is being implemented.[4] Note that $p_t(\gamma)$ can easily be computed from γ: If at period $t - 1$ action L has been chosen, then $p_t(\gamma) = r$. Namely, if by choosing L the object has not been found at period $t - 1$, then the object must have been at R at period $t - 1$, and hence will be at L with probability r next period. By a similar reasoning, if at period $t - 1$ action R has been chosen, then $p_t(\gamma) = 1 - q$. If at period $t - 1$ action W has been chosen, then

$$p_t(\gamma) = p_{t-1}(\gamma)(1 - q) + (1 - p_{t-1}(\gamma))r.$$

THEOREM 2.2: Let $P \geq 2$. Take a configuration $(\bar{p}, q, r)$ of initial and transition probabilities with $q \leq r$, let $\varepsilon > 0$, and choose the thresholds π_1 and π_2 according to Table 1. Then, the strategy γ given by

$$\gamma_t = \begin{cases} R, & \text{if } p_t(\gamma) \in [0, \pi_1] \\ W, & \text{if } p_t(\gamma) \in (\pi_1, \pi_2) \\ L, & \text{if } p_t(\gamma) \in [\pi_2, 1] \end{cases}$$

is an ε-optimal strategy.[5] Moreover, if $q = r$ or $q + r \geq 1$, then γ is an optimal strategy, unique up to the choices at π_1 and π_2.

In addition, γ is successful and prescribes to search a location only if the probability that the object is at that location is at least $\frac{1}{2}$. In particular, γ is loss-free.

Figure 2 provides a graphical illustration of the various areas of (q, r) pairs we distinguish in the theorem. Note that the assumption $q \leq r$ is made without loss of generality. If $q < r$ and $q + r \in (0, 1)$, then there are initial probabilities $\bar{p}$ for which no optimal strategy exists. However, we can still guarantee the existence of ε-optimal strategies in this case.

Consider the induced strategy γ in our main theorem. Regarding the number of consecutive periods that

---

[3] The reason is that at every period, if the object has not yet been found, the probabilities of the object being located at L and R would be exactly $\frac{1}{2}$. So, the expected searching cost will be $\frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \cdots = 2$.

[4] From now on, whenever we speak about period t, we always assume that the object has not been found before this period.

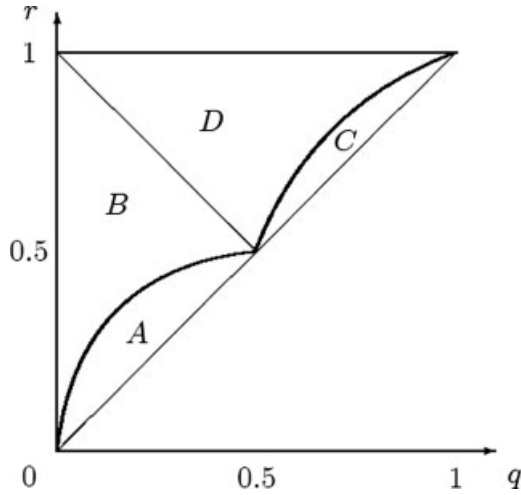[5] In fact, the strategy is ε-optimal from any period on.

Figure 2. Different areas of (q, r) pairs considered in Theorem 2.2.

γ will prescribe waiting, it will be clear from Section 3 that:

(1) in the case when $q + r \in (0, 1)$ and $p \in (\pi_1, \pi_2)$: for any $T \in \mathbb{N}$, if $\varepsilon > 0$ is small, then in certain situations, γ prescribes waiting for more than T consecutive periods;

(2) in the case when $q + r \in (1, 2)$ and $p \in (\pi_1, \pi_2)$: the strategy γ prescribes waiting at most once before searching a location.

It will follow from our analysis that our thresholds π_1 and π_2 have an interesting relationship with the threshold $\tilde{\pi}$ as computed by Pollock (1970) for the model without waiting: It turns out that $\tilde{\pi}$ is always in between π_1 and π_2. Consequently, if the agent searches a location in the model with waiting, he would search the same location in the model without waiting. Although this result may seem intuitive, it is not immediate.

The proof of Theorem 2.2 is structured as follows. Section 3 discusses all the different types of dynamics of the two-state Markov processes under consideration. Section 4 introduces the value function and analyzes its properties, which is fundamental for the proof of the main theorem. Section 5 contains preparatory lemmas for the proof of the main theorem. Section 6 contains two lemmas which prove the existence of the thresholds and show that the induced strategies are successful. We also show that these thresholds are unique if $q = r$ or $q + r \geq 1$. Consequently, we prove for these cases the uniqueness of the optimal strategy, up to the choices at the two thresholds. In Section 7, we compute the values of the thresholds π_1 and π_2 (see Table 1), and show that the induced strategy only searches a location if the object is there with probability at least $\frac{1}{2}$.

3. FIVE POSSIBLE TYPES OF DYNAMICS

We distinguish a number of different types of dynamics in the Markov process, induced by the transition probabilities q and r. To formally discuss these different types of dynamics, we need the following definitions.

Let p denote the probability that the object is at location L. If the agent waits, the probability that the object will be at L next period is given by

$$Ap := p(1 - q) + (1 - p)r = (1 - q - r)p + r. \quad (1)$$

By induction, the probability that the object will be at L after n periods of waiting is given by

$$A^n p := (1 - q - r)^n \left(p - \frac{r}{q + r} \right) + \frac{r}{q + r}, \quad (2)$$

unless $q + r = 0$.

Further, we denote the invariant distribution(s) of the Markov process by $(\pi^*, 1 - \pi^*)$. In particular, π^* is a solution of the equation

$$A\pi^* = \pi^* \Leftrightarrow (1 - q - r)\pi^* + r = \pi^*.$$

If $q + r > 0$, then this equation has a unique solution which is

$$\pi^* = \frac{r}{q + r}. \quad (3)$$

We distinguish the following types of dynamics.

(i) *Absorbing case:* $q + r = 0$, i.e. $q = r = 0$. Object does not move in this case. Function Ap is given by $Ap = p$. Invariant distribution π^* is not unique. In fact, any $\pi^* \in [0, 1]$ is invariant.

(ii) *Non-oscillating case:* $0 < q + r < 1$. Function Ap is strictly increasing in p. Invariant distribution π^* is unique. Probabilities Ap and p are always on the same side of π^*. That is, $A^n p$ converges to π^* in a monotonic (and hence non-oscillating) fashion.

(iii) *State independent transitions case:* $q + r = 1$. Function Ap is constant in p. Invariant distribution π^* is unique. The convergence to π^* is immediate. Since $r = 1 - q$, the transitions are independent of the state.

(iv) *Oscillating case:* $1 < q + r < 2$. Function Ap is strictly decreasing in p. Invariant distribution π^* is unique. Probabilities Ap and p are always on different sides of π^*. So, $A^n p$ converges to π^* in an oscillating fashion although $|A^n p - \pi^*|$ is a strictly decreasing function of $n \in \mathbb{N}$.

(v) *Switching case:* $q + r = 2$, i.e. $q = r = 1$. Function Ap is given by $Ap = 1 - p$. Invariant distribution

π^* is unique and equal to $\frac{1}{2}$. By waiting, the probability that the object is at L switches from p to $1-p$. So, unless $\bar{p} = \pi^*$, there will be no convergence to π^*.

4. VALUE FUNCTION

4.1. Definition

The key concept in the proof of our main theorem is that of a *value function*. Fix some transition probabilities q and r. Suppose that the object is located at L with probability p. Then, for every strategy γ we denote by $C(p, \gamma)$ the expected cost of finding the object if γ is used. The value $V(p)$ is given by

$$V(p) = \inf_{\gamma \, \text{successful}} C(p, \gamma)$$

and denotes the minimal expected cost of finding the object if the object is currently located at L with probability p. The function that assigns to every p the number $V(p)$ is called the value function. Hence, if a successful strategy γ is optimal for the configuration $(\bar{p}, q, r)$, then $C(\bar{p}, \gamma) = V(\bar{p})$. Conversely, if $P \geq 2$ and γ is successful with $C(\bar{p}, \gamma) = V(\bar{p})$, then, in view of Lemma 2.1, γ is optimal.

Note that $C(p, \gamma) = pC(1, \gamma) + (1 - p)C(0, \gamma)$ is linear in p. Hence, $V(p)$, being the infimum of linear functions of p, is a concave function of p.

4.2. Basic Properties of the Value Function

For each action $a \in \{L, R, W\}$, we define the expected cost $V(p, a)$ induced by using action a, given that the agent behaves optimally afterwards. As before, p denotes the probability that the object is currently at L.

When the agent searches L, then with probability p he finds the object and his total cost will be 1, while with probability $(1 - p)$ he cannot find it. In case he cannot find it, the object was at location R, and the probability of it being at L next period is r. Hence, he will have a future expected cost of $V(r)$ if he acts optimally from next period on. Therefore, the expected cost of searching L can be written as

$$V(p, L) = p \cdot 1 + (1 - p)[1 + V(r)] = 1 + (1 - p)V(r). \tag{4}$$

Similarly, the expected cost of searching R is

$$V(p, R) = (1 - p) \cdot 1 + p[1 + V(1 - q)]$$
$$= 1 + pV(1 - q). \tag{5}$$

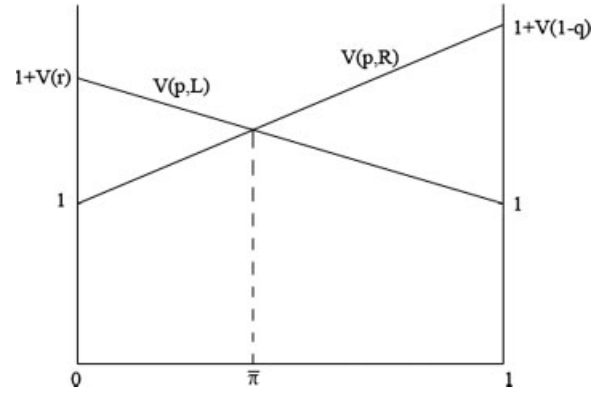Finally, the expected cost of waiting is

$$V(p, W) = V(Ap). \tag{6}$$

Figure 3. Functions $V(p, L)$ and $V(p, R)$.

Recall that Ap denotes the probability that the object will be at L in the next period, if the object is currently at L with probability p.

Therefore, the value function satisfies

$$V(p) = \min\{V(p, L), V(p, R), V(p, W)\}. \tag{7}$$

An action $a \in \{L, R, W\}$ with $V(p, a) = V(p)$ is called an optimal action at p. Moreover, for a given $\varepsilon > 0$ we say that action a is ε-optimal at p if $V(p, a) \leq V(p) + \varepsilon$.

As it can be seen from the equations above, $V(p, L)$ and $V(p, R)$ are linear in p. Moreover, $V(p, L)$ is strictly monotone decreasing in p, whereas $V(p, R)$ is strictly monotone increasing in p. Combining this with the fact that $V(p, R)$ attains its minimum, which is equal to 1, at $p = 0$ and $V(p, L)$ attains its minimum, which is equal to 1, at $p = 1$, we can conclude that $V(p, L)$ and $V(p, R)$ always intersect at a unique $\bar{\pi} \in (0, 1)$, which is given by the equation

$$\bar{\pi} = \frac{V(r)}{V(1 - q) + V(r)}. \tag{8}$$

Hence,

$$V(p, L) > V(p, R) \quad \text{for all } p \in [0, \bar{\pi})$$
$$V(p, L) = V(p, R) \quad \text{at } p = \bar{\pi}$$
$$V(p, L) < V(p, R) \quad \text{for all } p \in (\bar{\pi}, 1].$$

Figure 3 provides a graphical representation of the functions $V(p, L)$ and $V(p, R)$.

Recall that $V(p)$ is a concave function of p. Since Ap is linear in p, it follows that $V(p, W) = V(Ap)$ is also a concave function of p. Note that $V(A0) = V(r)$ and $V(A1) = V(1 - q)$. As we will argue in the lemma below, both $V(r)$ and $V(1 - q)$ are greater than or equal to 1. Therefore, the concave function $V(Ap)$ can cross each of the functions $V(p, L)$ and $V(p, R)$ at most once. From (7),

it follows that the optimal action is given by at most two thresholds.

In the following lemma we list some general properties of the value function that will be useful for the proof of the main theorem.

LEMMA 4.1: The value function $V(p)$ satisfies the following properties;

  (i) $1 \leq V(p) \leq 2$, and $V(p) = 1 \Leftrightarrow p = 0$ or $p = 1$.
  (ii) $V(p) = \inf_{n \in \{0,1,...\}} \min\{V(A^n p, L), V(A^n p, R)\}$.
  (iii) $V(\pi^*) = \min\{V(\pi^*, L), V(\pi^*, R)\}$ if $q + r > 0$.
  (iv) $V(p) \leq V(A^n p) \ \forall n \in N$, and $V(p) \leq V(\pi^*)$ if $0 < q + r < 2$.
  (v) $V(1 - q) \leq \frac{1}{q}$ if $q > 0$ and $V(r) \leq \frac{1}{r}$ if $r > 0$.

PROOF: (i) It is clear that $V(p) \geq 1$ and $V(p) = 1 \Leftrightarrow p = 0$ or $p = 1$. It remains to prove that $V(p) \leq 2$. Consider the strategy γ which, at every period t, searches the location that contains the object with the highest probability (if both locations are equally likely, check L). Then, by the proof of Lemma 2.1, we know that $C(p, \gamma) \leq 2$, yielding $V(p) \leq 2$.

(ii) Since $V(p)$ restricts to successful strategies, the agent can only wait for a finite number of periods in a row, after which L or R has to be chosen.

(iii) This follows from (ii) and the fact that $A^n \pi^* = \pi^*$ for every n.

(iv) By definition, $V(p) \leq V(p, W) = V(Ap)$. Applying transformation A recursively, we get

$$V(p) \leq V(Ap) \leq V(A^2 p) \leq \cdots \leq V(A^{n-1} p) \leq V(A^n p)$$

for every n.

Since $0 < q + r < 2$, the sequence $A^n p$ converges to π^*. By (ii), $V(p)$ is continuous in p, and hence we have that $V(A^n p)$ converges to $V(\pi^*)$. Therefore, $V(p) \leq \lim_{n \to \infty} V(A^n p) = V(\pi^*)$.

(v) We have $V(1-q) \leq V(1-q, R) = 1 + (1-q)V(1-q)$. Hence, if $q > 0$, this implies

$$V(1 - q) \leq \frac{1}{q}.$$

Similarly, one can show that $V(r) \leq \frac{1}{r}$, if $r > 0$. □

5. PREPARATORY LEMMAS

To prove our main theorem, we first show that a sequence of optimal actions, under mild conditions, always leads to an optimal strategy.

LEMMA 5.1: Let γ be a strategy which never waits infinitely long. Let $p_t(\gamma)$ be the probability that at period t the

object is located at L if the object has not been found yet, and strategy γ is being implemented. Suppose that there is an $\alpha > 0$ such that L is chosen only if $p_t(\gamma) \geq \alpha$, and R is chosen only if $1 - p_t(\gamma) \geq \alpha$. Then, γ is successful and

  (1) if every action γ_t is optimal at $p_t(\gamma)$ then strategy γ is optimal;
  (2) for every $\varepsilon > 0$ and $\delta \in [0, \alpha\varepsilon]$, if every action $\gamma_t \in \{L, R\}$ is δ-optimal at $p_t(\gamma)$ and every action $\gamma_t = W$ is optimal at $p_t(\gamma)$, then strategy γ is ε-optimal.

PROOF: Since γ never waits infinitely long, it can be described as follows: First, wait until period w_1, then search location $a_1 \in \{L, R\}$, then wait until period w_2, then search location a_2, and so on. For $k = 1, 2, 3, \ldots$ define

$$z_k := \begin{cases} p_{w_k}(\gamma), & \text{if } a_k = L \\ 1 - p_{w_k}(\gamma), & \text{if } a_k = R. \end{cases}$$

Hence, z_k denotes the probability that the object will be found by action a_k at period w_k, if the object has not been found yet, and strategy γ is being implemented. Note that, by our assumption in the lemma, $z_k \geq \alpha$ for all k. From now on, we will write p_t instead of $p_t(\gamma)$.

We now show that γ is successful. The probability that γ never finds the object is $\Pi_{k \geq 1}(1 - z_k)$. Since $z_k \geq \alpha$ for all k, we have that $1 - z_k \leq 1 - \alpha < 1$. Hence, $\Pi_{k \geq 1}(1 - z_k) = 0$, yielding that γ is successful.

We now prove (1) and (2) by showing that for every $\varepsilon \geq 0$, if $\delta \in [0, \alpha\varepsilon]$, every action $\gamma_t \in \{L, R\}$ is δ-optimal at $p_t(\gamma)$, and every action $\gamma_t = W$ is optimal at $p_t(\gamma)$, then strategy γ is ε-optimal. Part (1) would then follow by taking $\varepsilon = 0$, and part (2) would follow by taking $\varepsilon > 0$.

Choose $\varepsilon \geq 0$ and $\delta \in [0, \alpha\varepsilon]$. Suppose now that at every period t every action $\gamma_t \in \{L, R\}$ is δ-optimal for $p_t(\gamma)$, and every action $\gamma_t = W$ is optimal at $p_t(\gamma)$. We will show that γ is an ε-optimal strategy, i.e. $C(\bar{p}, \gamma) \leq V(\bar{p}) + \varepsilon$, where $\bar{p}$ is the initial probability that the object is at L.

For every period t, let γ^t be the continuation strategy from period t onwards. Then,

$$C(\bar{p}, \gamma) = z_1 \cdot 1 + (1 - z_1)\left[1 + C(p_{w_1+1}, \gamma^{w_1+1})\right],$$

since the object will be found at period w_1 with probability z_1, while with probability $1 - z_1$ the continuation strategy γ^{w_1+1} would be played after period w_1. On the other hand,

$$V(\bar{p}) = V(p_{w_1}) \geq V(p_{w_1}, a_1) - \delta$$
$$= z_1 \cdot 1 + (1 - z_1)[1 + V(p_{w_1+1})] - \delta.$$

Here, the first equality follows from the fact that waiting until period w_1 is a sequence of optimal actions, and the

inequality follows from the assumption that a_1 is a δ-optimal action at period w_1. The last equality follows from (4) and (5). Therefore,

$$V(\bar{p}) - C(\bar{p}, \gamma)$$
$$\geq (1 - z_1)\left[V(p_{w_1+1}) - C(p_{w_1+1}, \gamma^{w_1+1})\right] - \delta,$$

and by induction it would follow that

$$V(\bar{p}) - C(\bar{p}, \gamma) \geq (1 - z_1) \cdot \cdots \cdot (1 - z_k)$$
$$\times \left[V(p_{w_k+1}) - C(p_{w_k+1}, \gamma^{w_k+1})\right]$$
$$- \delta \left[1 + (1 - z_1) + \cdots + (1 - z_1) \cdots (1 - z_{k-1})\right] \quad (9)$$

for $k = 1, 2, 3 \ldots$

We will show that $V(\bar{p}) - C(\bar{p}, \gamma) \geq -\varepsilon$, which would imply that γ is an ε-optimal strategy. We already know that $\Pi_{k \geq 1}(1 - z_k) = 0$. Hence, it is sufficient to show that

(i) $V(p_{w_k+1}) - C(p_{w_k+1}, \gamma^{w_k+1})$ is uniformly bounded in k, and
(ii) $\delta[1 + (1 - z_1) + \cdots + (1 - z_1) \cdots (1 - z_{k-1})] \leq \varepsilon$ for all k.

To show (i), since by Lemma 4.1 (i), $1 \leq V(p_{w_k+1}) \leq 2$, it remains to prove that $C(p_{w_k+1}, \gamma^{w_k+1})$ is uniformly bounded. As $z_k \geq \alpha$ for all k, it holds that

$$C(p_{w_k+1}, \gamma^{w_k+1}) \leq \alpha \cdot 1 + (1 - \alpha) \cdot \alpha \cdot 2$$
$$+ (1 - \alpha)^2 \cdot \alpha \cdot 3 + \cdots = \frac{1}{\alpha},$$

and hence $C(p_{w_k+1}, \gamma^{w_k+1})$ is uniformly bounded by $\frac{1}{\alpha}$.

To show (ii), for every k,

$$\delta[1 + (1 - z_1) + \cdots + (1 - z_1) \cdots (1 - z_{k-1})]$$
$$\leq \alpha \varepsilon [1 + (1 - \alpha) + \cdots + (1 - \alpha)^{k-1}] \leq \alpha \varepsilon \frac{1}{\alpha} = \varepsilon.$$

This completes the proof of the lemma. $\qquad \square$

The next result relates the transition probabilities, q and r, to the values of π^* and $\bar{\pi}$, which will be useful for the proof of the main theorem.

LEMMA 5.2: For any transition probabilities $q, r \in [0, 1]$ with $q + r > 0$, it holds that (i) $q < r \iff \bar{\pi} < \pi^*$, (ii) $q > r \iff \bar{\pi} > \pi^*$ and (iii) $q = r \iff \bar{\pi} = \pi^*$.

PROOF: It is sufficient to show that the following three implications are valid: (i) $q < r \implies \bar{\pi} < \pi^*$, (ii) $q > r \implies \bar{\pi} > \pi^*$, (iii) $q = r \implies \bar{\pi} = \pi^*$.

Proof of (i). Suppose $q < r$. In view of (8) and (3), $\bar{\pi} < \pi^*$ is equivalent to

$$qV(r) < rV(1 - q). \quad (10)$$

Thus, we have to show that (10) holds. If $V(r) < V(1 - q)$, then (10) is obvious. So suppose $V(r) > V(1 - q)$. This means that $r \neq 1 - q$, which together with the assumption $q < r$ leaves us only two cases: $q + r \in (0, 1)$ (the non-oscillating case) and $q + r \in (1, 2)$ (the oscillating case).

Assume $q + r \in (0, 1)$ (the non-oscillating case). By (3) we have $r < \pi^* < 1 - q$. Since $A^n r$ is monotonically converging to π^*, we have for all $n \in \mathbb{N}$, $V(r, L) > V(A^n r, L) > V(A^{n+1}, L)$ and $\lim_{n \to \infty} V(A^n r, L) = V(\pi^*, L)$, as well as $V(r, R) < V(A^n r, R)$. Hence by property (ii) of Lemma 4.1 we obtain

$$V(r) = \min\{V(\pi^*, L), V(r, R)\}.$$

Similarly,

$$V(1 - q) = \min\{V(1 - q, L), V(\pi^*, R)\}$$

Since $V(\pi^*, R) > V(r, R)$, by our assumption $V(r) > V(1 - q)$, inequality $V(1 - q, L) < V(r, R)$ follows. Hence, by (4) and (5),

$$1 + qV(r) = V(1 - q, L) < V(r, R) = 1 + rV(1 - q)$$

yielding (10).

Assume now $q + r \in (1, 2)$ (the oscillating case). By (3) we have $1 - q < \pi^* < r$. Due to the facts that $A(1 - q) > (1 - q)$ and $A(1 - q) > A^n(1 - q)$ for all $n > 1$ we can conclude that $V(A(1 - q), L) < V(1 - q, L)$ and $V(A(1 - q), L) < V(A^n(1 - q), L)$ for all $n > 1$.

Also by $(1 - q) < A^n(1 - q)$ for all $n \in \mathbb{N}$ we have $V(1 - q, R) < V(A^n(1 - q), R)$ for all $n \in \mathbb{N}$. Therefore, by property (ii) of Lemma 4.1, $V(1 - q) = \min\{V(A(1 - q), L), V(1 - q, R)\}$ follows.

Similarly to the arguments above, $V(r) = \min\{V(r, L), V(Ar, R)\}$. Since, by (1), $A(1 - q) < r$, we obtain $V(A(1 - q), L) > V(r, L)$. Hence, by $V(r) > V(1 - q)$, we must have $V(1 - q, R) < V(r, L)$ yielding $V(1 - q) = V(1 - q, R)$. Therefore,

$$V(1 - q) = V(1 - q, R) = 1 + (1 - q)V(1 - q)$$

From the assumption that $q + r \in (1, 2)$, it follows that $q > 0$. Therefore, from the equation above, $V(1 - q) = \frac{1}{q}$. In view of property (v) of Lemma 4.1, we may then conclude that $qV(r) \leq \frac{q}{r} < \frac{r}{q} = rV(1 - q)$, which implies (10).

Proof of (ii). Similar to the proof of (i) above.

Proof of (iii). Assume that $q = r$. Then, as transition probabilities q and r are equal, the problem is symmetric, and $V(p) = V(1 - p)$ for all $p \in [0, 1]$. Hence, by (8) and (3), we obtain $\bar{\pi} = \frac{1}{2}$ and $\pi^* = \frac{1}{2}$, hence $\bar{\pi} = \pi^*$ follows. $\qquad \square$

6. EXISTENCE OF THRESHOLDS

In this section we will show that we can always find thresholds π_1 and π_2 such that the induced strategy γ is (ε-)optimal and successful. We also show that these thresholds are unique if $q = r$ or $q + r \geq 1$. Consequently, we prove for these cases the uniqueness of the optimal strategy, up to the choices at the two thresholds. The exact computation of the thresholds follows in Section 7.

In view of Lemma 5.1, it is sufficient to prove the following two lemmas.

LEMMA 6.1: Take a configuration $(\bar{p}, q, r)$ of initial and transition probabilities with $q \leq r$. Then, there are thresholds $\pi_1, \pi_2 \in [0, 1]$ with $\pi_1 \leq \pi_2$, $\pi_1 < 1$ and $\pi_2 > 0$ such that: searching R is an optimal action for every $p \in [0, \pi_1]$, waiting is an optimal action for every $p \in (\pi_1, \pi_2)$, and searching L is an optimal action for every $p \in [\pi_2, 1]$. Here, the optimal actions are unique up to the choices at π_1 and π_2. As usual, p denotes the probability for the object being located at L.

Moreover, if $q = r$ or $q + r \geq 1$, then the induced strategy γ never waits infinitely long.

Also, if $q + r > 0$, then $\pi_2 = \pi^*$.

We will prove this lemma later in Sections 6.1.–6.5.

LEMMA 6.2: Take a configuration $(\bar{p}, q, r)$ of initial and transition probabilities such that $0 < q + r < 1$ (i.e., the non-oscillating case) and $q < r$. Then, there is a threshold $\pi_1 \in [0, \pi^*]$ satisfying $\pi_1 < 1$ with the following property: For every $\delta > 0$ there is a $\tau > 0$ such that searching R is an optimal action for every $p \in [0, \pi_1]$, waiting is an optimal action for every $p \in (\pi_1, \pi^* - \tau)$, and searching L is a δ-optimal action for every $p \in [\pi^* - \tau, 1]$. Moreover, the induced strategy γ never waits infinitely long.

We will prove this lemma in Section 6.6.
We will now prove Lemma 6.1. We distinguish five cases.

6.1. Absorbing Case: $q + r = 0$

The fact that $q = r = 0$ implies directly that the object does not move. Hence, waiting makes no sense. By searching a location at period 1, one either finds the object immediately, or one will be sure to find the object next period at the other location. Therefore, only L is optimal if $p > \frac{1}{2}$, only R is optimal if $p < \frac{1}{2}$, and both are optimal at $p = \frac{1}{2}$. Accordingly, $\pi_1 = \pi_2 = \bar{\pi} = \frac{1}{2}$. Note that the optimal actions are unique up to the choice at $\frac{1}{2}$, and moreover the induced strategy never waits.

6.2. Nonoscillating Case: $0 < q + r < 1$

Since $q \leq r$, by Lemma 5.2 we have that $\bar{\pi} \leq \pi^*$. Define $\pi_2 = \pi^*$ and let $\pi_1 \in [0, \bar{\pi}]$ be the unique p for which $V(p, R) = V(\pi^*, L)$. Note that π_1 is well-defined for the following reasons: First, $V(0, R) = 1 \leq V(\pi^*, L)$ due to Lemma 4.1, (i). Also, $V(\bar{\pi}, R) = V(\bar{\pi}, L) \geq V(\pi^*, L)$ due to the definition of $\bar{\pi}$, monotonicity of $V(p, L)$, and the fact that $\pi^* \geq \bar{\pi}$. Hence, there exists a $p \in [0, \bar{\pi}]$ with $V(p, R) = V(\pi^*, L)$. Moreover, since $V(p, R)$ is strictly increasing, there can only be one such p.

Note that $\pi_2 > 0$ as $\pi_2 = \pi^* = \frac{r}{q+r}$, $r \geq q$ and $q + r > 0$. Also, as $\bar{\pi} < 1$ we have that $\pi_1 < 1$.

On the interval $[0, \pi_1)$: It holds that $V(p, R) < V(p, L)$ since $\pi_1 \leq \bar{\pi}$. Since $p < A^n p$ for all n, we have that $V(p, R) < V(A^n p, R)$ for all n. Moreover, since $A^n p < \pi^*$, it holds that $V(p, R) < V(\pi_1, R) = V(\pi^*, L) \leq V(A^n p, L)$. Therefore, by Lemma 4.1, (ii), $V(p) = V(p, R)$, and only R is optimal at $[0, \pi_1)$.

It similarly follows that R is optimal at π_1.

On the interval (π_1, π_2): We have that $V(p, R) > V(\pi_1, R) = V(\pi^*, L) \geq V(\pi^*)$, and $V(p, L) > V(\pi_2, L) = V(\pi^*, L) \geq V(\pi^*)$. Since, by Lemma 4.1, (iv), $V(p) \leq V(\pi^*)$, it follows that $V(p) = V(p, W)$, and therefore only waiting is optimal at (π_1, π_2).

On the interval $(\pi_2, 1]$: We know that $V(p, L) < V(p, R)$. Since $A^n p < p$ for all n, we have that $V(p, L) < V(A^n p, L)$ for all n. Moreover, since $A^n p > \pi^*$, it holds that $V(p, L) < V(A^n p, L) \leq V(A^n p, R)$. Therefore, by Lemma 4.1, (ii), $V(p) = V(p, L)$, and only L is optimal at $(\pi_2, 1]$.

It similarly follows that L is optimal at π_2.

So, the optimal actions are unique up to the choice at π_1 and π_2.

Note that if $q = r$, then $\pi^* = \bar{\pi}$ by Lemma 5.2, hence $\pi_1 = \pi_2 = \pi^*$. Therefore, the induced strategy never waits.

6.3. State Independent Transitions Case: $q + r = 1$

The proof for this case is identical to the previous case.

6.4. Oscillating Case: $1 < q + r < 2$

Since $q \leq r$, by Lemma 5.2 we have that $\bar{\pi} \leq \pi^*$. Define $\pi_2 = \pi^*$ and let $\pi_1 \in [0, \bar{\pi}]$ be the unique p for which $V(p, R) = V(Ap, L)$. Note that π_1 is well-defined for the following reasons: First, $V(0, R) = 1 \leq V(A0, L)$ due to the Lemma 4.1, (i). Also, since $\pi^* \geq \bar{\pi}$, we have $V(\bar{\pi}, R) = V(\bar{\pi}, L) \geq V(\pi^*, L) \geq V(A\bar{\pi}, L)$, where the last inequality follows from the fact that $A\bar{\pi} \geq \pi^*$. Hence, there exists a $p \in [0, \bar{\pi}]$ with $V(p, R) = V(Ap, L)$. Now, suppose that more than one $p \in [0, 1]$ would exist with $V(p, R) = V(Ap, L)$. Since both sides are linear functions in p, it must be the case that $V(p, R) = V(Ap, L)$ for all

$p \in [0, 1]$. In particular, $1 = V(0, R) = V(A0, L)$, which by Lemma 4.1, (i) implies that $A0 = 1$, and hence $r = 1$. Also, $V(\bar{\pi}, L) = V(\bar{\pi}, R) = V(A\bar{\pi}, L)$, which implies that $\bar{\pi} = A\bar{\pi}$, and hence $\bar{\pi} = \pi^*$. It follows from Lemma 5.2 that $q = r$, and hence $q = r = 1$, which contradicts $q + r < 2$. So, there can only be one p with $V(p, R) = V(Ap, L)$.

Note that $\pi_2 > 0$ as $\pi_2 = \pi^* = \frac{r}{q+r}, r \geq q$ and $q + r > 1$. Also, as $\bar{\pi} < 1$ we have that $\pi_1 < 1$.

It also follows that

$$V(p, R) < V(Ap, L) \text{ for all } p \in [0, \pi_1), \text{ and} \quad (11)$$

$$V(p, R) > V(Ap, L) \text{ for all } p \in (\pi_1, 1]. \quad (12)$$

On the interval $[0, \pi_1)$ it holds that $V(p, R) < V(p, L)$ since $\pi_1 \leq \bar{\pi}$. Since $A^n p > p$ for all n, we have that $V(p, R) < V(A^n p, R)$ for all n. Moreover, by (11), $V(p, R) < V(Ap, L) \leq V(A^n p, L)$, where the latter follows from the fact that $A^n p \leq Ap$ for all n. Therefore, by Lemma 4.1, (ii), $V(p) = V(p, R)$, and therefore only R is optimal at $[0, \pi_1)$.

It similarly follows that R is optimal at π_1.

On the interval (π_1, π_2), we have that $V(Ap, L) < V(p, L)$ since $Ap > p$. Moreover, $V(Ap, L) < V(p, R)$ by (12). In view of $V(p, W) = V(Ap) \leq V(Ap, L)$, it follows that only waiting is an optimal action at (π_1, π_2).

On the interval $(\pi_2, 1]$, we know that $V(p, L) < V(p, R)$. Since $p > A^n p$ for all n, we have that $V(p, L) < V(A^n p, L)$ for all n. Since $\pi^* \geq \bar{\pi}$, we have that $V(\pi^*, L) \leq V(\pi^*, R) = V(A\pi^*, R)$. Also, $V(1, L) = 1 < V(A1, R)$ as $A1 = 1 - q > 0$. Since $V(p, L)$ and $V(Ap, R)$ are linear in p, it follows that $V(p, L) < V(Ap, R)$ on $(\pi_2, 1]$. Moreover, $V(Ap, R) \leq V(A^n p, R)$ since $A^n p \geq Ap$ for all n. Therefore, $V(p, L) < V(A^n p, R)$ for all n. Together with $V(p, L) < V(A^n p, L)$ for all n, it follows from Lemma 4.1, (ii), that $V(p) = V(p, L)$, and only L is optimal at $(\pi_2, 1]$.

It similarly follows that L is optimal at π_2.

So, the optimal actions are unique up to the choice at π_1 and π_2, and the induced strategy waits at most once.

6.5. Switching Case: $q + r = 2$

In this case, $q = r = 1$, so the object will surely move to the other location if it is not found. By searching a location at period 1, one either finds the object immediately, or one will be sure to find the object next period at the same location. Therefore, only L is optimal if $p > \frac{1}{2}$, only R is optimal if $p < \frac{1}{2}$, and both are optimal at $p = \frac{1}{2}$. Accordingly, $\pi_1 = \pi_2 = \bar{\pi} = \frac{1}{2}$. Note that the optimal actions are unique up to the choice at $\frac{1}{2}$, and moreover the induced strategy never waits.

6.6. Near Optimality in the Nonoscillating Case: $0 < q + r < 1$ and $q < r$

We now prove Lemma 6.2. Choose π_1 as in Lemma 6.1. Take a $\delta > 0$, and choose $\tau > 0$ small enough such that $\pi^* - \tau \geq \bar{\pi}$ and $V(\pi^* - \tau, L) - V(\pi^*, L) \leq \delta$. Such a τ exists due to Lemma 5.2 (i), and the continuity of $V(p, L)$. On the interval $[\pi^* - \tau, \pi^*]$ we have that (1) $V(p, R) \geq V(p, L)$, (2) $A^n p$ converges to π^* in a monotonically increasing way, and (3) $V(p, L)$ is decreasing in p. Therefore, by Lemma 4.1 (ii), $V(p) = V(\pi^*, L)$ for every $p \in [\pi^* - \tau, \pi^*]$. Now, take some $p \in [\pi^* - \tau, \pi^*]$. Then,

$$V(p, L) - V(p) = V(p, L) - V(\pi^*, L)$$
$$\leq V(\pi^* - \tau, L) - V(\pi^*, L) \leq \delta,$$

which means that searching L is a δ-optimal action on $[\pi^* - \tau, \pi^*]$. As on the interval $[\pi^*, 1]$, searching L is an optimal action (cf. Section 6.2), we obtain that searching L is a δ-optimal action on the whole $[\pi^* - \tau, 1]$. It is clear that the strategy induced by π_1 and $\pi^* - \tau$ never waits infinitely long.

7. COMPUTATION OF THRESHOLDS

In this section we derive exact formulas for the thresholds π_1 and π_2 in Lemmas 6.1 and 6.2. In Section 7.1 we deal with Lemma 6.1, and also show that $\pi_1 \leq \frac{1}{2}$ and $\pi_2 \geq \frac{1}{2}$, which will imply that the induced strategies only prescribe to search a location if the probability that this location contains the object is at least $\frac{1}{2}$. In Section 7.2 we deal with the computation for Lemma 6.2.

7.1. Computation for Lemma 6.1

Note that for the cases $q + r = 0$ and $q + r = 2$ we have already shown that $\pi_1 = \pi_2 = \frac{1}{2}$ (see Sections 6.1. and 6.5.). So, it remains to analyze the cases where $0 < q + r < 1$, $q + r = 1$, and $1 < q + r < 2$. Recall our assumption that $q \leq r$. By Lemma 5.2 we know that $\bar{\pi} \leq \pi^*$. Recall that $\pi_2 = \pi^* = \frac{r}{q+r} \geq \frac{1}{2}$. So, we only need to compute π_1 and show that $\pi_1 \leq \frac{1}{2}$.

7.1.1. Nonoscillating Case: $0 < q + r < 1$

Since $q + r < 1$, we have $\pi^* > r$, and $1 - \pi^* = \frac{q}{q+r} > q$, which implies $\pi^* < 1 - q$. So,

$$r < \pi^* < 1 - q.$$

We distinguish two cases here: (i) $r \leq \pi_1$ and (ii) $r > \pi_1$. We will show that case (i) corresponds to the case in Table 1

where $(q+1)r^2 + q^2 r - q \leq 0$, and that case (ii) corresponds to the case where $(q+1)r^2 + q^2 r - q > 0$.

CASE 1: $r \leq \pi_1$.

By Lemma 6.1 we know that

$$V(r) = V(r, R), \text{ and}$$
$$V(1-q) = V(1-q, L) \text{ since } 1 - q > \pi^* = \pi_2.$$

From equations (4) and (5), it follows

$$V(r) = 1 + rV(1-q) \text{ and } V(1-q) = 1 + qV(r),$$

which implies that

$$V(r) = \frac{1+r}{1-rq} \text{ and } V(1-q) = \frac{1+q}{1-rq}.$$

Then, by equations (4) and (5) we can write

$$V(p, R) = 1 + p\frac{1+q}{1-rq} \text{ and}$$
$$V(p, L) = 1 + (1-p)\frac{1+r}{1-rq}.$$

Recall that π_1 is the unique solution to $V(\pi_1, R) = V(\pi^*, L)$. Since

$$V(\pi_1, R) = 1 + \pi_1 \frac{1+q}{1-rq} \text{ and}$$
$$V(\pi^*, L) = 1 + \left(1 - \frac{r}{q+r}\right)\frac{1+r}{1-rq},$$

we get

$$\pi_1 = \frac{q(1+r)}{(q+r)(1+q)}.$$

An easy calculation shows that $r \leq \pi_1$ if and only if $(q+1)r^2 + q^2 r - q \leq 0$. Also, it may be verified that $q \leq r$ implies $\pi_1 \leq \frac{1}{2}$.

CASE 2: $r > \pi_1$.

Since $r < \pi^*$, we know from Lemma 6.1 that

$$V(r) = V(r, W) = V(Ar).$$

Since $Ar \in (\pi_1, \pi^*)$, we obtain by induction that

$$V(r) = V(A^n r)$$

for all n, and hence

$$V(r) = \lim_{n \to \infty} V(A^n r) = V(\pi^*).$$

As $\pi^* \geq \bar{\pi}$, we have by Lemma 4.1, (iii) that $V(\pi^*) = V(\pi^*, L)$. Hence,

$$V(r) = V(\pi^*, L).$$

Since $1 - q > \pi^* = \pi_2$, Lemma 6.1 yields $V(1-q) = V(1-q, L)$.

Combining these insights with equations (4) and (5) we obtain

$$V(1-q) = 1 + qV(r) \text{ and } V(r) = 1 + (1-\pi^*)V(r).$$

From these two equations we get

$$V(r) = \frac{q+r}{r} \text{ and } V(1-q) = 1 + q\frac{q+r}{r}.$$

Recall that π_1 is the unique solution to $V(\pi_1, R) = V(\pi^*, L)$. Since

$$V(\pi_1, R) = 1 + \pi_1 V(1-q) \text{ and}$$
$$V(\pi^*, L) = V(r) = \frac{q+r}{r},$$

we obtain that

$$\pi_1 = \frac{q}{(q+r)(1+q) - q}.$$

An easy calculation shows that $r > \pi_1$ if and only if $(q+1)r^2 + q^2 r - q > 0$.

We will now show that $\pi_1 \leq \frac{1}{2}$. If $r \leq \frac{1}{2}$, then $\pi_1 < r$ yields $\pi_1 \leq \frac{1}{2}$. Suppose now that $r > \frac{1}{2}$. Let $\phi(a, b) := a^2 + a(b-2) + b$ for all real numbers a and b. Then, the inequality $\pi_1 \leq \frac{1}{2}$ is equivalent with $\phi(q, r) \geq 0$. Notice that, given any b, the parabola $\phi(a, b)$ is minimal at $a^* = 1 - \frac{b}{2}$. By using that

$$q < 1 - r \leq 1 - \frac{r}{2},$$

we obtain

$$\phi(q, r) > \phi(1 - r, r) = -1 + 2r \geq 0,$$

since $r > \frac{1}{2}$. Hence, $\pi_1 \leq \frac{1}{2}$.

7.1.2. State Independent Transitions Case: $q + r = 1$

In this case, $\pi_2 = \pi^* = r = 1 - q$. Since $q \leq r$, we know that $q \leq \frac{1}{2}$. We show that $\pi_1 = q$, which will imply that $\pi_1 \leq \frac{1}{2}$. By copying the arguments above, we obtain the following:

CASE 1: $r \leq \pi_1$.

We find that

$$\pi_1 = \frac{q(1+r)}{(q+r)(1+q)} = \frac{q(1+r)}{(1+q)}.$$

We note that $r \leq \pi_1$ and $q \leq r$ yield $q = r$, and hence $\pi_1 = q$.

CASE 2: $r > \pi_1$.

We get

$$\pi_1 = \frac{q}{(q+r)(1+q) - q} = q.$$

7.1.3. Oscillating Case: $1 < q + r < 2$

Since $q + r > 1$, we have $\pi^* < r$, and $1 - \pi^* = \frac{q}{q+r} < q$, which implies $\pi^* > 1 - q$. So,

$$1 - q < \pi^* < r.$$

We distinguish two cases here: (i) $1 - q \leq \pi_1$ and (ii) $1 - q > \pi_1$. We will show that case (i) corresponds to the case in Table 1 where $(1 - q + q^2)r - 2q^2 + q^3 \leq 0$, while case (ii) corresponds to the case where $(1 - q + q^2)r - 2q^2 + q^3 > 0$.

CASE 1: $1 - q \leq \pi_1$.

Since $\pi^* = \pi_2$, we know by Lemma 6.1 that

$$V(1 - q) = V(1 - q, R) \text{ and } V(r) = V(r, L).$$

From Eqs. (4) and (5), it follows

$$V(1 - q) = 1 + (1 - q)V(1 - q) \text{ and}$$
$$V(r) = 1 + (1 - r)V(r).$$

Hence

$$V(1 - q) = \frac{1}{q} \text{ and } V(r) = \frac{1}{r}.$$

Then, by Eqs. (4) and (5) we can write

$$V(p, R) = 1 + p\frac{1}{q} \text{ and } V(p, L) = 1 + (1 - p)\frac{1}{r}.$$

Recall that π_1 is the unique solution to $V(\pi_1, R) = V(A\pi_1, L)$. Since $A\pi_1 = (1 - q - r)\pi_1 + r$, we have

$$V(\pi_1, R) = 1 + \pi_1\frac{1}{q} \text{ and}$$
$$V(A\pi_1, L) = 1 + (1 - (1 - q - r)\pi_1 - r)\frac{1}{r}.$$

This yields

$$\pi_1 = \frac{q(1 - r)}{(q + r)(1 - q)}.$$

An easy calculation shows that $1 - q \leq \pi_1$ if and only if $(1 - q + q^2)r - 2q^2 + q^3 \leq 0$. It can also be verified that $q \leq r$ implies $\pi_1 \leq \frac{1}{2}$.

CASE 2: $1 - q > \pi_1$.

Since $\pi^* = \pi_2$, we know by Lemma 6.1 that

$$V(1 - q) = V(1 - q, W) \text{ and } V(r) = V(r, L).$$

As $V(1 - q, W) = V(A(1 - q))$ and $A(1 - q) > \pi^* = \pi_2$, we obtain $V(1 - q, W) = V(A(1 - q), L)$. Since $A(1 - q) = (1 - q - r)(1 - q) + r$, it follows from equations (4) and (5) that

$$V(1 - q) = 1 + (1 - (1 - q - r)(1 - q) - r)V(r)$$
$$\text{and } V(r) = 1 + (1 - r)V(r).$$

Hence

$$V(r) = \frac{1}{r} \text{ and } V(1 - q) = \frac{q + (q + r)(1 - q)}{r}.$$

Then, by Eqs. (4) and (5) we can write

$$V(p, R) = 1 + p\frac{q + (q + r)(1 - q)}{r} \text{ and}$$
$$V(p, L) = 1 + (1 - p)\frac{1}{r}.$$

Recall that π_1 is the unique solution to $V(\pi_1, R) = V(A\pi_1, L)$. Since $A\pi_1 = (1 - q - r)\pi_1 + r$, we have

$$V(\pi_1, R) = 1 + \pi_1\frac{q + (q + r)(1 - q)}{r} \text{ and}$$
$$V(A\pi_1, L) = 1 + (1 - (1 - q - r)\pi_1 - r)\frac{1}{r}.$$

This yields

$$\pi_1 = \frac{1 - r}{(q + r)(1 - q) + 1 - r}.$$

An easy calculation shows that $1 - q > \pi_1$ if and only if $(1 - q + q^2)r - 2q^2 + q^3 > 0$.

We will now show that $\pi_1 \leq \frac{1}{2}$. As $q + r > 1$ and $q \leq r$, we have $r > \frac{1}{2}$. Let $\phi(a, b) := -a^2 + a(1 - b) + (2b - 1)$ for all real numbers a and b. Then, the inequality $\pi_1 \leq \frac{1}{2}$ is equivalent with $\phi(q, r) \geq 0$. Notice that, given any b, the parabola $\phi(a, b)$ is maximal at $a^* = \frac{1-b}{2}$. By using that

$$\frac{1 - r}{2} < 1 - r < q \leq r$$

we obtain

$$\phi(q, r) > \phi(r, r) = -2r^2 + 3r - 1 = (1 - r)(2r - 1) \geq 0,$$

since $r > \frac{1}{2}$. Hence, $\pi_1 \leq \frac{1}{2}$.

7.2. Computation for Lemma 6.2

We now compute the threshold $\pi^* - \tau$ in Lemma 6.2. Recall that $0 < q + r < 1$ and $q < r$. Take a $\delta > 0$, and choose $\tau := \min\{\pi^* - \frac{1}{2}, \frac{\delta}{2}\}$. Note that $\tau > 0$ and $\pi^* - \tau \geq \frac{1}{2}$. We know, from Section 7.1.1., that $\pi_1 \leq \frac{1}{2}$, and hence $\pi^* - \tau \geq \frac{1}{2} \geq \pi_1$. Therefore, on the interval $[\pi^* - \tau, \pi^*]$ we have that (1) $V(p, R) \geq V(\pi^*, L)$, (2) $V(p, L) \geq V(\pi^*, L)$, and (3) $A^n p$ converges to π^* in a monotonically increasing way. Hence, by Lemma 4.1 (ii), $V(p) = V(\pi^*, L)$ for every $p \in [\pi^* - \tau, \pi^*]$. Now, take some $p \in [\pi^* - \tau, \pi^*]$. In order to prove that searching L is a δ-optimal action, it is sufficient to show that

$$V(p, L) \leq V(p) + \delta = V(\pi^*, L) + \delta.$$

We know, by (4), that

$$V(p, L) = 1 + (1 - p)V(r)$$

and

$$V(\pi^*, L) = 1 + (1 - \pi^*)V(r).$$

Hence $V(p, L) - V(\pi^*, L) = (\pi^* - p)V(r)$. But $(\pi^* - p) \leq \frac{\delta}{2}$ by the choice of p. Moreover, $V(r) \leq 2$ by Lemma 4.1, (i). Therefore,

$$V(p, L) - V(\pi^*, L) = (\pi^* - p)V(r) \leq \delta.$$

8. RELATION TO THE MODEL WITHOUT WAITING

In this section we compare the optimal strategy we determined in the previous sections with the optimal strategy that Pollock [6] found for the model without waiting. We thus evaluate the precise consequence of introducing the option to wait. Recall that the optimal strategy in Pollock [6] is given by only one threshold $\tilde{\pi}$, meaning that the agent searches R if $p < \tilde{\pi}$ and searches L if $p > \tilde{\pi}$, and is indifferent if $p = \tilde{\pi}$. We note that Pollock's threshold $\tilde{\pi}$ is always in between our thresholds π_1 and π_2, which is far from being trivial, and is shown in Flesch et al. [2]. Consequently, if the agent searches a location in the model with waiting, he would search the same location in the model without waiting.

Below, we will illustrate, by means of an example from Ross [7], that in the model without waiting the agent may search a location that contains the object with probability less than one half (i.e. $\tilde{\pi} \neq \frac{1}{2}$). Moreover, we will compare the values obtained in the two models.

8.1. Ross' Example Revisited

An important implication of our model is that it is never optimal to search the location with lower current probability of containing the object (See Theorem 2.2). In the model without the option to wait, as Ross [7] has shown, it is possible that searching the location with the lower probability of containing the object is optimal (see also [9]). In fact, this is a striking result of the model without waiting. This happens if searching the location with the lower probability of containing the object serves as an investment for the future, i.e. decreases the uncertainty about the location of the object in the future. The following example by Ross [7] shows this.

Example [7] $\bar{p} = 0.45$ and $1 - \bar{p} = 0.55$, $q = \frac{1}{2}$ and $r = 1$.

Consider first the model without waiting. Although the initial probability of location R containing the object is higher, it is optimal to search L since searching L gives complete certainty about the location of the object at $t = 2$ in case it is not found at $t = 1$. One can see that, in contrast to searching L, searching R at $t = 1$ leads to complete uncertainty about the location of the object at $t = 2$, in case it is not found at $t = 1$. As a result, the expected cost induced by searching L at $t = 1$ and acting optimally afterwards (i.e., search L again) is 1.55 and the expected cost induced by searching R at $t = 1$ and acting optimally afterwards (i.e., search L, and if the object is not found then search L once more) is 1.675. In fact, it follows from Pollock's calculations that $\tilde{\pi} = 0.4 < \frac{1}{2}$.

Let's now introduce the option to wait. Since $1 < q + r < 2$ and $(1 - q + q^2)r - 2q^2 + q^3 > 0$ in this example, we see from Table 1 that the thresholds are $\pi_1 = 0$ and $\pi_2 = \frac{2}{3}$. Notice the large difference with the model without waiting. Hence, at period 1 the agent would wait instead of searching L. After waiting, the probability that the object is located at L is $(0.45)(0.5) + (0.55)1 = 0.775 > \pi_2$. Hence, the agent would search L at period 2. If at period 2, he does not find the object, he will surely find it in period 3 by searching L. Accordingly, the expected cost of finding the object is $(0.775) \cdot 1 + (0.225) \cdot 2 = 1.225$, which is less than the expected searching cost without the option to wait (i.e., 1.55).

8.2. Comparison of Values

We finally compare the value $\tilde{V}$ in the model without waiting with the value V in the model with waiting. Clearly, $\tilde{V} \geq V$ always.

We will first investigate the situation where the difference between the two values is minimal. It is intuitive that when $q = r = 0$ (Case 1) or $q = r = 1$ (Case 5), we have $\pi_1 = \pi_2 = \tilde{\pi} = \frac{1}{2}$ and hence $\tilde{V} = V$. In those cases, namely, there is no reason for the agent to wait. Hence, optimal strategies in the model with waiting and the optimal strategies in the model without waiting coincide.

Now, consider the situation where $q = 0$, $r = 1$ (a special case of Case 3). In this case, the agent would like to wait whenever $p \in (0, 1)$, so the waiting region is maximal here. Now suppose that the initial probability is $\bar{p} = \frac{1}{2}$. Then, $\tilde{V}(\frac{1}{2}) = \frac{3}{2}$, since by searching a location at $t = 1$ the agent will find the object immediately with probability $\frac{1}{2}$, and otherwise will find the object for sure at location L next period. On the other hand, in the model with waiting we have $V(\frac{1}{2}) = 1$, since by waiting at $t = 1$ the agent will find the object for sure at location L next period. We conjecture that the difference between the two values can never be more than $\frac{1}{2}$, so that this would be a case where the difference is maximal.

9. DISCUSSION

1. Possibility of overlooking. Weber [10] and MacPhee and Jordan [4] analyzed the search for a moving target problem with a single agent and without waiting in continuous time and discrete time, respectively. Their models take the possibility of overlooking into account. Thus, in their models, even if the object is at location L (or R) and the agent searches that location, he may overlook the object with some positive probability. This possibility drastically affects the updating process. Now, when the agent searches a location and cannot find the object, he should take into account the possibility that the object was in fact at the location he searched. Hence, it is not clear from our analysis that the two-threshold form will hold for the (nearly) optimal strategy if overlooking is possible.

2. Different searching costs for left and right. It can be simply shown that our results do not change qualitatively since $V(p, L)$ and $V(p, R)$ will remain to be linear functions of p and $V(Ap)$ will still be a concave function of p. The (nearly) optimal strategy for this case can still be described by a two-threshold form.

3. Costly Waiting. We assumed that waiting is costless, i.e., $c_W = 0$. Now, assume that $c_W > 0$. In this case, $V(p, W) = c_W + V(Ap)$ will still be a concave function of p. Moreover, the arguments in Section 4 that the optimal actions can be described by at most two thresholds are still valid. However, we can now prove the existence of an optimal strategy even in the case $0 < q + r < 1$, in contrast to the case $c_W = 0$. The reason is that waiting is not an optimal action at π^* and hence neither of the thresholds is equal to π^*.

4. Existence of optimal strategies. In the model without the option to wait, optimal strategies always exist

(cf. [4, 6]). However, as we mentioned in the analysis above, when waiting is a (costless) option, an optimal strategy may not exist in the non-oscillating case (i.e., $0 < q + r < 1$). This is because, for certain initial probabilities, waiting one more period always provides a more favorable probability distribution over the two locations. Nevertheless, we proved that for every $\varepsilon > 0$, there exists an ε-optimal strategy.

5. More than two locations. The case of more than two locations is much more difficult than the case of two locations. Even in the case of three locations, one should provide two-dimensional areas of probability distributions that describe an optimal search strategy instead of thresholds on a line. (See Nakai [5], MacPhee and Jordan [4] and Thomas and Eagle [8] for some results for the model with more than two locations.)

REFERENCES

[1] D. Assaf and A. Sharlin-Bilitzky, Dynamic search for a moving target, J Appl Probab 31 (1994), 438–457.

[2] J. Flesch, E. Karagözoğlu, and A. Perea, "Optimal search for a moving target with the option to wait," Reseach Memorandum #51, Maastricht University, 2007.

[3] Y.C. Kan, Optimal search of a moving target, Operat Res 25 (1977), 864–870.

[4] I.M. MacPhee and B.P. Jordan, Optimal search for a moving target, Probab Eng Inform Sci 9 (1995), 159–182.

[5] T. Nakai, A model of search for a target among three boxes: some special cases, J Operat Res Soc Jpn 16 (1973), 151–162.

[6] S.M. Pollock, A simple model of search for a moving target, Operat Res 18 (1970), 883–903.

[7] S. Ross, Introduction to stochastic dynamic programming, Academic Press, New York, 1983.

[8] L.C. Thomas and J.N. Eagle, Criteria and approximate methods for path-constrained moving-target search problems, Nav Res Logistics 42 (1995), 27–38.

[9] A.R. Washburn, On a search for a moving target, Nav Res Logistics 27 (1980), 315–322.

[10] R.R. Weber, Optimal search for a randomly moving object, J Appl Probab 23 (1986), 708–717.

[11] D.J. White, A conjecture of Ross for a search problem, J Math Anal Appl 166 (1992), 263–271.