European Journal of Operational Research 58 (1992) 78–89
North-Holland

## Theory and Methodology

# Asymptotic analysis of dichotomous search with search and travel costs

Refael Hassin and Reuven Hotovely
*Department of Statistics, Tel Aviv University, Tel Aviv 69978, Israel*

Received June 1990

**Abstract:** An object to be searched lies on $[0, n]$. The search starts from the left end of the interval. A query at a point finds out whether the objects lies to the left or to the right of the point. The problem is to determine a sequence of queries that minimizes the expected sum of query ($c_Q$ per attempt) and travel ($c_T$ per unit distance) costs required to locate the object within an interval of unit length. This work develops simple approximate solutions to the problem and analyzes their performance.

**Keywords:** Dichotomous search, dynamic programming

## 1. Introduction

An underground communication line is found to be cut off. A technician wants to locate a segment of the line, of unit length, that contains the cut off point, and replace it. Assuming that the line is cut off in exactly one point, it follows that from each point on it, it is possible to communicate with exactly one of its ends. This provides a mechanism for testing whether the cut off point is 'before' or 'after' the checked point. Two costs are involved with the search; a travel cost, proportional to the traveled distance, and a query cost, proportional to the number of points in which a test is performed. The technician has no prior knowledge about the cut off point, and his objective is to locate it, within a segment of unit length, with minimum expected cost.

The problem can be put in a more abstract form: The location of an object to be searched is a random variable uniformly distributed on $(0, n]$. The search starts from an end of the interval. A query at a point finds out whether the object lies to the left or to the right of the point. In order to place a query at a point the searcher must first travel to that point. The problem is to determine a sequence of queries that minimizes the expected sum of query ($c_Q$ per attempt) and travel ($c_T$ per unit distance) costs required to locate the object within an interval of unit length.

While it is a simple task to write down a recursive equation that describes the process, an explicit form of the policy does not seem to exist. This work analyzes several simple approximations to the optimal policy. Section 2 deals with *fixed-step policies*, where the searcher advances by a fixed distance as long as the direction of his movement is fixed. Section 3 and the Appendix deal with *fixed-ratio policies*. A fixed-ratio policy is characterized by a number $p \in (0, 1)$. Given an interval of size $n > 1$ that contains the object, the next query is placed at a distance of $np$ from the end point in which the searcher is located. Section 4 deals with *myopic policies* intended to maximize the reduction in the interval's length per unit of cost associated with the first query. For each class of approximations we single out a preferred member. All of the resulting policies are asymptotically optimal. Section 5 contains numerical comparisons for a set of interval lengths.

In many applications queries are restricted to a finite set of points, and in particular to $n$ points that are equally spaced over the interval. This restriction is not significant when large intervals are considered. The approximations we present naturally fit into the unrestricted model. Nevertheless, we will present them assuming that queries are restricted to integer points. This especially adds to the complexity of analyzing fixed-ratio policies, and the quality of continuous approximations to the travel and query costs of the restricted case are presented in the appendix.

Our main results are that a step of size $\sqrt{nc_Q/c_T}$ is approximately optimal for the fixed-step class, and yields a cost of approximately $\frac{1}{2}c_T n + \sqrt{nc_Q c_T}$. The same order of cost is produced by a ratio of $2\sqrt{c_Q/nc_T}$ (that is, a first step of $2\sqrt{nc_Q/c_T}$), which is approximately optimal for this class. A first step of $\sqrt{nc_Q/c_T}$ defines approximately also an optimal myopic policy with a cost of order $c_T n + 2\sqrt{nc_Q c_T}$. However, in this case it is possible to refine the approximation to obtain better solutions for short intervals than in the other cases. The relative error of all of the above approximations tends to 0 at least as fast as $1/\sqrt{n}$.

There is a vast literature concerning dichotomous search problems, mostly aimed at computing optimal alphabetic trees (see, for example, Gilbert and Moore, 1959, Hu and Tucker, 1971). Closest to the present subject is Murakami's paper (1976) where an identical situation is analyzed but with the objective of minimizing the maximum cost (over all possible locations of the object) required to locate the object. Asymptotic results for a related problem were obtained by Cameron and Narayanamurthy (1964) and by Murakami (1971). There, it is assumed that the cost of a query placed to the right of the object is different from that of a query placed to its left and an asymptotically optimal fixed-ratio policy is derived. However, no travel costs are considered in these papers. A different approach to obtain approximate policies is described by Hassin (1984) where the a priori distribution of the object's location is assumed to geometric. The approximation there, which is quite accurate also for small intervals, is achieved by relaxing an integer programming formulation of the problem.

**Definition 1.1.** A policy is a rule that determines the place of the first query (with respect to the end containing the searcher) for any given length of interval. A policy is optimal if it minimizes the expected sum of query and travel costs.

**Definition 1.2.** Let $g(n)$ be an approximation for $f(n)$. The relative error of this approximation is $|(f(n)-g(n)/f(n)|$.

We denote by $F_\pi(n)$ the expected cost associated with searching an interval of length $n$ while using policy $\pi$ at each decision point of the search.

Let $F(n)$ be the expected cost associated with an optimal policy. Then $F(1)=0$, and for $n=2,3,\ldots$:

$$F(n) = \min_{x=1,\ldots,n-1} F(n,x)$$

where

$$F(n,x) = c_Q + c_T x + \frac{x}{n}F(x) + \frac{n-x}{n}F(n-x). \tag{1.1}$$

**Definition 1.3.** A policy $\pi$ is asymptotically optimal if

$$\lim_{n\to\infty} \frac{F_\pi(n)}{F(n)} = 1.$$

In the rest of this paper we mean $\log_2$ whenever we refer to log. We use $o(f(n))$ to denote a function $g(n)$ satisfying $\lim_{n\to\infty}(g(n)/f(n))=0$.

## 2. Fixed-step policies

Let us start our analysis by mentioning two extreme policies. The first is the *unit-step search* where, as long as $n > 1$, the next query is placed at a distance of one unit from the location of the searcher (that is, $x = 1$). The expected cost of this policy is $(c_Q + c_T)(1 + 2 + \cdots + (n-1) + (n-1))/n = (c_Q + c_T)((n+1)/2 - 1/n)$. This policy is clearly optimal when $c_Q$ is very small relative to $c_T$, and it minimizes the expected travel distance. The other extreme is the *binary search* where the interval is divided into approximately two equal parts. To be exact, suppose that $n = 2^r + \rho$ where $r$ and $\rho$ are nonnegative integers satisfying $\rho < 2^r$. Then $x$ is set to $2^{r-1}$ if $\rho \leqslant 2^{r-1}$ and to $\rho$ otherwise. This policy is the unique optimal one when $c_T$ is very small relative to $c_Q$ but is still positive. It minimizes the expected number of queries and has the least travel distance among all such policies. The expected cost of this policy is $c_Q \log n + c_T(n-1)$ when $\rho = 0$. (See Morris, 1969, and Theorem 7 of Hassin, 1984, for exact derivation of the query cost when $\rho > 0$ and a discussion of alternative optimal policies when $c_T = 0$.) The following lemma results from the above observations.

**Lemma 2.1.** *For* $n = 2, 3, \ldots,$

$$F(n) \geqslant \tfrac{1}{2} c_T n + c_Q \log n.$$

**Proof.** The first term on the right-hand side is a lower bound on the expected travel cost for any policy, while the second is a lower bound on the expected query cost. □

**Corollary 2.2.** *A policy* $\pi$ *such that* $F_\pi(n) = \frac{1}{2} c_T n + o(n)$ *is asymptotically optimal.*

The next theorem is intended to demonstrate that it is easy to produce asymptotically optimal policies.

**Theorem 2.3.** *Each policy in the following class is asymptotically optimal: Let* $0 < \alpha \leqslant \beta < 1$ *be fixed. For* $n = 2, 3, \ldots,$ *let* $x$ *satisfy* $n^\alpha \leqslant x \leqslant n^\beta$. *Place each query at a distance* $\lfloor x \rfloor$ *to your right as long as the object is to your right and the remaining interval is strictly longer than* $\lfloor x \rfloor$. *Else, continue the search using unit-step search.*

**Proof.** Let $\pi$ be any policy of the type described in the theorem. The expected cost of unit-search applied to an interval of size at most $x$ is less than $x(c_T + c_Q)$. The expected travel distance till the unit-search is applied is at most $\frac{1}{2}n + n^\beta$. Therefore,

$$F_\pi(n) \leqslant c_Q \frac{n}{n^\alpha} + c_T\left(\tfrac{1}{2}n + n^\beta\right) + (c_Q + c_T)n^\beta = c_Q n^{1-\alpha} + \tfrac{1}{2} c_T n + (c_Q + 2c_T) n^\beta. \tag{2.1}$$

By Corollary 2.2 the policy is asymptotically optimal. □

From the theorem we conclude that to compare approximate policies one should also consider secondary terms in the cost function that affect the rate of convergence.

Considering the bound on $F_\pi(n)$ in (2.1), we notice that except for the dominating term $\frac{1}{2} c_T n$ there are two secondary terms proportional to $n^{1-\alpha}$ and $n^\beta$. To obtain fast convergence of $F_\pi(n)$ to $F(n)$ we would like to minimize the sum of these terms. This is obtained by minimizing $\max\{1-\alpha, \beta \mid 0 < \alpha \leqslant \beta\}$, that is, by setting $\alpha = \beta = 0.5$. Thus the best policy in the class of Theorem 2.3 places the next query at a point whose distance from the end containing the searcher is of order $\sqrt{n}$. This defines the following *fixed-step policy*: By placing queries at distance $\lfloor \gamma\sqrt{n} \rfloor$ locate an interval of this size or less. At this time change the step size according to the size of the present interval and repeat the process for this interval. The expected cost of this policy, $h(n)$, satisfies

$$h(n) = \tfrac{1}{2} c_T n + c_Q \tfrac{1}{2} \frac{n}{\gamma\sqrt{n}} + h\left(\lfloor \gamma\sqrt{n} \rfloor\right) + o(\sqrt{n}) = \tfrac{1}{2} c_T n + \tfrac{1}{2}\left(\frac{c_Q}{\gamma} + \gamma c_T\right)\sqrt{n} + o(\sqrt{n}).$$

The middle term is minimized when we take $\gamma = \sqrt{c_Q/c_T}$ and then

$$x = \sqrt{n\frac{c_Q}{c_T}}, \tag{2.2}$$

and

$$h(n) = \tfrac{1}{2}c_T n + \sqrt{nc_Q c_T} + o(\sqrt{n}). \tag{2.3}$$

From (2.3) and Lemma 2.1 we obtain:

**Corollary 2.4.** *The relative error of the fixed-step policy with a step as in* (2.2) *converges to* 0 *as fast as* $1/\sqrt{n}$.

## 3. Fixed-ratio policies

It is easily seen from (1.1) that for any $x > n/2$, $F(n, x) \geqslant F(n, n-x)$. Therefore there exists an optimal policy with $x \leqslant n/2$. In this section we approximate the optimal policy by a fixed-ratio policy, characterized by a constant $0 < p < 1$ so that the next query is placed at $x = \max\{1, \lfloor np \rfloor\}$. From the above observation, it suffices to consider values $0 < p \leqslant 0.5$. We make this assumption throughout this section whenever we refer to $p$.

The expected travel distance under a fixed-ratio policy with parameter $p$ is recursively given by

$$D(1, p) = 0, \qquad D(2, p) = 1,$$

and for $n = 3, 4, \ldots,$

$$D(n, p) = \lfloor np \rfloor + \frac{\lfloor np \rfloor}{n} D(\lfloor np \rfloor, p) + \frac{n - \lfloor np \rfloor}{n} D(n - \lfloor np \rfloor, p), \quad p \geqslant 1/n, \tag{3.1a}$$

$$D(n, p) = 1 + \frac{n-1}{n} D(n-1, p) = \frac{n+1}{2} - \frac{1}{n}, \qquad p < 1/n. \tag{3.1b}$$

Let

$$d(y, p) = \frac{y-1}{2(1-p)}. \tag{3.2}$$

The following lemma shows that $d$ has properties that are similar to those of $D$, except that the floor operation is missing in its functional equation.

**Lemma 3.1.** $d(1, p) = 0,$

$$d(y, p) = yp + pd(yp, p) + (1-p)d(y(1-p), p).$$

**Proof.**

$$\begin{aligned} &yp + pd(yp, p) + (1-p)d(y(1-p), p) \\ &= yp + \frac{p(yp-1)}{2(1-p)} + \frac{(1-p)[y(1-p)-1]}{2(1-p)} \\ &= \frac{1}{2(1-p)}\left[yp2(1-p) + yp^2 - p + y(1-p)^2 - (1-p)\right] = \frac{y-1}{2(1-p)}. \qquad \square \end{aligned}$$

The proof of the following theorem is presented in the Appendix:

**Theorem 3.2.** *For* $n = 2, 3, \ldots,$

$$|d(n, p) - D(n, p)| < 2 \log n.$$

Since $D(n, p) \geq \frac{1}{2}n$, the following corollary obtains:

**Corollary 3.3.**

$$\lim_{n \to \infty} \frac{d(n, p)}{D(n, p)} = 1.$$

*The relative error converges to* 0 *at least as fast as* $(\log n)/n$.

The expected number of queries under a fixed-ratio policy is given by

$$S(1, p) = 0, \qquad S(2, p) = 1,$$

and for $n = 3, 4, \ldots,$

$$S(n, p) = 1 + \frac{\lfloor np \rfloor}{n} S(\lfloor np \rfloor, p) + \frac{n - \lfloor np \rfloor}{n} S(n - \lfloor np \rfloor, p), \quad p \geq 1/n, \tag{3.3a}$$

$$S(n, p) = 1 + \frac{n-1}{n} S(n-1, p) = \frac{n+1}{2} - \frac{1}{n}, \quad p < 1/n. \tag{3.3b}$$

Let $H(p)$ be the *binary entropy* of $p$ (see, for example, Mehlhorn, 1984):

$$H(p) = -p \log p - (1-p) \log(1-p).$$

**Remark 3.4.** $H(p)$ is positive and monotone increasing on (0, 0.5]; $H(0.5) = 1$.

For $y > 0$ define

$$s(y, p) = \frac{\log y}{H(p)}.$$

The following lemma shows that $s$ has properties that are similar to those of $S$, except that the floor operation is missing in its functional equation.

**Lemma 3.5.** $s(1, p) = 0$, *and*

$$s(y, p) = 1 + ps(yp, p) + (1-p)s(y(1-p), p).$$

**Proof.**

$$s(y, p) = 1 + ps(yp, p) + (1-p)s(y(1-p), p) = 1 + \frac{p \log yp}{H(p)} + \frac{(1-p) \log y(1-p)}{H(p)}$$

$$= 1 + \frac{p \log p + (1-p) \log(1-p)}{H(p)} + \frac{\log y}{H(p)} = 1 - 1 + \frac{\log y}{H(p)} = s(y, p). \qquad \square$$

The proof of the following theorem is presented in the Appendix.

**Theorem 3.6.** *Suppose* $p < \frac{1}{3}$. *Then*

$$|s(n, p) - S(n, p)| < \frac{1}{pH(p)} \log \frac{4}{p}.$$

Since $S(n, p) \geq \log n$, the following corollary obtains for $p < \frac{1}{3}$:

**Corollary 3.7.**

$$\lim_{n\to\infty} \frac{s(n,\, p)}{S(n,\, p)} = 1.$$

*The relative error converges to 0 at least as fast as* $1/\log n$.

Let

$$\hat{F}(n) = \min_p \hat{F}(n,\, p)$$

where

$$\hat{F}(n,\, p) = c_Q s(n,\, p) + c_T d(n,\, p) = \frac{c_Q \log n}{H(p)} + \frac{c_T(n-1)}{2(1-p)}.$$

Differentiating $\hat{F}(n,\, p)$ with respect to $p$ and equating to 0 we obtain

$$c_Q \log n \frac{\log((1-p)/p)}{[H(p)]^2} = c_T \frac{n-1}{2(1-p)^2}$$

or,

$$\frac{(n-1)c_T}{2c_Q \log n} = (1-p)^2 \frac{\log(1-p) - \log p}{[-p \log p - (1-p) \log(1-p)]^2}. \tag{3.4}$$

Consider now the case where $c_Q$ and $c_T$ are kept constant, while $n \to \infty$. The left-hand side of the above expression tends to infinity, and thus for equality to hold the right-hand side must tend to infinity as well. Since $p \in (0, 0, 5]$ this means that $p \to 0$. In this case, the right-hand side can be approximated by

$$\frac{-\log p}{p^2 \log^2 p} = \frac{-1}{p^2 \log p}.$$

Let $q = 1/p$, then we would like to approximate a solution to

$$\frac{c_T n}{2c_Q \log n} = \frac{q^2}{\log q}.$$

Substituting $q = k\sqrt{n}$ we obtain

$$\frac{c_T n}{2c_Q \log n} = \frac{k^2 n}{\log k + 0.5 \log n}.$$

For large $n$ we can approximate $k$ by $k = \frac{1}{2}\sqrt{c_T/c_Q}$. This means

$$p = \frac{1}{q} = \frac{1}{k\sqrt{n}} = \frac{2}{\sqrt{n}}\sqrt{\frac{c_Q}{c_T}}.$$

Substituting $p$ in $\hat{F}(n,\, p)$ we obtain

$$\hat{F}(n,\, p) \approx \frac{c_Q \log n}{-(2/\sqrt{n})\sqrt{c_Q/c_T} \log\left(2/\sqrt{n}\sqrt{c_Q/c_T}\right)} + \frac{c_T n}{2} \approx \sqrt{n c_Q c_T} + c_T \frac{n}{2}. \tag{3.5}$$

**Theorem 3.8.** *The fixed-ratio policy with* $p = (2/\sqrt{n})\sqrt{c_Q/c_T}$ *is asymptotically optimal. Its cost,* $g(n)$, *satisfies*

$$g(n) = \tfrac{1}{2} c_T n + \sqrt{n c_T c_Q} + o(\sqrt{n}). \tag{3.6}$$

**Proof.**

$$g(n) = c_Q S(n, p) + c_T D(n, p)$$
$$= \hat{F}(n, p) + c_Q[S(n, p) - s(n, p)] + c_T[D(n, p) - d(n, p)].$$

By Theorem 3.6,

$$| s(n, p) - S(n, p) | \leqslant \frac{1}{pH(p)} \log \frac{4}{p}.$$

Substituting $p$ we obtain that $|s(n, p) - S(n, p)| = O(1)$. The claim now follows from (3.5), Lemma 2.1, and Theorem 3.2. □

## 4. Myopic policies

Recall that

$$F(n) = \min_{x=1,\ldots,n-1} F(n, x)$$

where

$$F(n, x) = c_Q + c_T x + \frac{x}{n} F(x) + \frac{n-x}{n} F(n-x).$$

This recurrence can be given the following interpretation: By investing $c_Q + x c_T$ we reduce the size of the interval to $x$ with probability $x/n$, and to $n - x$ with probability $(n-x)/n$. A myopic policy is one that maximizes the expected reduction in the size of the interval per unit of cost. The expected reduction is

$$\frac{x}{n}(n-x) + \frac{n-x}{n} x = \frac{2x(n-x)}{n}.$$

Thus we want to solve

$$\max_x \frac{2}{n} \frac{x(n-x)}{c_Q + x c_T}. \tag{4.1}$$

Differentiating and equating to 0 we obtain a quadratic equation whose nonnegative solution is

$$x = -\frac{c_Q}{c_T} + \sqrt{\left(\frac{c_Q}{c_T}\right)^2 + n\left(\frac{c_Q}{c_T}\right)}. \tag{4.2}$$

For large $n$, $x$ is approximated by

$$x = \sqrt{n \frac{c_Q}{c_T}}. \tag{4.3}$$

From (4.1) the cost per unit reduction in the length of the interval known to contain the object is

$$\frac{n}{2} \frac{c_Q + x c_T}{x(n-x)} = \tfrac{1}{2} c_T \left(1 + \frac{2c_Q}{\sqrt{n c_Q c_T} - c_Q}\right).$$

Denote by $G(n)$ the cost under the myopic policy with $x = \lfloor \sqrt{n c_Q / c_T} \rfloor$, in a problem of size $n$, then

$$G(n) \approx \tfrac{1}{2} n c_T + c_Q c_T \sum_{k=1}^{n} \frac{1}{\sqrt{k c_Q c_T} - c_Q}.$$

Table 1
$F_{fs} / F_{opt}$

| $c_Q / c_T$ | $n$ | | | | |
|---|---|---|---|---|---|
| | 100 | 200 | 500 | 600 | 800 |
| 800/1 | 1.0020 | 1.0035 | 1.0013 | 1.0078 | 1.0110 |
| 500/1 | 1.0031 | 1.0056 | 1.0047 | 1.0128 | 1.0184 |
| 300/1 | 1.0051 | 1.0090 | 1.0143 | 1.0235 | 1.0340 |
| 100/1 | 1.0141 | 1.0283 | 1.0523 | 1.0553 | 1.0614 |
| 3/1 | 1.0612 | 1.0528 | 1.0350 | 1.0335 | 1.0298 |
| 1/1 | 1.0446 | 1.0276 | 1.0194 | 1.0181 | 1.0158 |
| 1/3 | 1.0135 | 1.0145 | 1.0100 | 1.0103 | 1.0085 |
| 1/100 | 1.0000 | 1.0000 | 1.0018 | 1.0012 | 1.0007 |
| 1/300 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1/500 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1/800 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

For large $n$,

$$G(n) \approx \tfrac{1}{2} n c_T + \sqrt{c_Q c_T} \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \approx \tfrac{1}{2} n c_T + 2\sqrt{n c_Q c_T}.$$

It is interesting to note that under the asymptotic myopic solution of (4.3) (and the fixed-step policy of (2.2)) the first step is half its size under the fixed-ratio policy described in Theorem 3.7. This can be explained by noting that in the myopic policy the ratio increases while $n$ decreases. The fixed-ratio policy cannot increase the ratio along the search process while smaller intervals are considered, and therefore it starts from the beginning with a larger ratio.

## 5. Numerical results

Let $F_{opt}$ denote the expected cost under an optimal policy.

Let $F_{fs}$ denote the expected cost under a policy with a step of size $\max\{1, \min\{\lfloor x \rfloor, \lfloor \frac{1}{2} n \rfloor\}\}$, where $x$ is the step defined by the fixed-step policy of (2.2).

Let $F_{fr}$ denote the expected cost under a fixed-ratio policy with a step of size $\max\{1, \min\{\lfloor np \rfloor, \lfloor \frac{1}{2} n \rfloor\}\}$, where $p$ is as in Theorem 3.8.

Let $F_m$ denote the expected cost under the myopic policy with a step of $\max\{1, \min\{\lfloor x \rfloor, \lfloor \frac{1}{2} n \rfloor\}\}$, where $x$ is given by (4.3). We note that using (4.2) rather than (4.3) did not yield a significant improvement for the data presented in this section.

Tables 1–3 give some numerical examples for the ratios of the expected costs of the approximate to the optimal policies. We see that the convergence of this ratio to 1 is faster as the ratio $c_Q/c_T$ decreases. We also notice that for large ratios $c_Q/c_T$, all of the three policies give very close values. As this ratio decreases, the fixed-ratio policy becomes inferior. For very small ratios all the three policies are very close to optimal.

## Appendix

**Lemma A.1.** *For $n \geqslant 5$,*

$$\log(n - \lfloor np \rfloor) + p < \log n.$$

Table 2
$F_{fr} / F_{opt}$

| $c_Q / c_T$ | $n$ | | | | |
|---|---|---|---|---|---|
| | 100 | 200 | 500 | 600 | 800 |
| 800/1 | 1.0020 | 1.0035 | 1.0013 | 1.0078 | 1.0110 |
| 500/1 | 1.0031 | 1.0056 | 1.0047 | 1.0128 | 1.0184 |
| 300/1 | 1.0051 | 1.0090 | 1.0143 | 1.0235 | 1.0340 |
| 100/1 | 1.0141 | 1.0283 | 1.0680 | 1.0860 | 1.1136 |
| 3/1 | 1.1138 | 1.0770 | 1.0495 | 1.0467 | 1.0418 |
| 1/1 | 1.0545 | 1.0398 | 1.0271 | 1.0258 | 1.0231 |
| 1/3 | 1.0239 | 1.0163 | 1.0107 | 1.0100 | 1.0087 |
| 1/100 | 1.0004 | 1.0020 | 1.0011 | 1.0009 | 1.0008 |
| 1/300 | 1.0000 | 1.0000 | 1.0007 | 1.0006 | 1.0006 |
| 1/500 | 1.0000 | 1.0000 | 1.0000 | 1.0003 | 1.0004 |
| 1/800 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

**Proof.** Define $f(p) = 1 - p - 2^{-p}$. Then $f$ is monotone decreasing for $p > 0$, and $f(0.5) < -0.2$. Therefore,

$$\frac{n - \lfloor np \rfloor}{n} - 2^{-p} < \frac{n - np + 1}{n} - 2^{-p} = f(p) + \frac{1}{n} < 0$$

for $n \geqslant 5$. This is equivalent to $(n - \lfloor np \rfloor)2^{p} < n$. Taking the logarithm on both sides yields the lemma's claim. □

Let $A(n, p) = d(n, p) - D(n, p)$.

**Lemma A.2.** *For* $n < 1/p$,

$$|A(n, p)| < 1.$$

**Proof.** From (3.1), for $n < 1/p$

$$A(n, p) = \tfrac{1}{2}\frac{n-1}{1-p} - \tfrac{1}{2}(n+1) + \frac{1}{n}.$$

Table 3
$F_{m} / F_{opt}$

| $c_Q / c_T$ | $n$ | | | | |
|---|---|---|---|---|---|
| | 100 | 200 | 500 | 600 | 800 |
| 800/1 | 1.0020 | 1.0035 | 1.0013 | 1.0078 | 1.0110 |
| 500/1 | 1.0031 | 1.0056 | 1.0047 | 1.0128 | 1.0184 |
| 300/1 | 1.0051 | 1.0090 | 1.0143 | 1.0235 | 1.0340 |
| 100/1 | 1.0141 | 1.0283 | 1.0523 | 1.0558 | 1.0520 |
| 3/1 | 1.0265 | 1.0151 | 1.0078 | 1.0066 | 1.0050 |
| 1/1 | 1.0083 | 1.0038 | 1.0019 | 1.0017 | 1.0013 |
| 1/3 | 1.0000 | 1.0001 | 1.0000 | 1.0001 | 1.0000 |
| 1/100 | 1.0000 | 1.0000 | 1.0002 | 1.0001 | 1.0001 |
| 1/300 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1/500 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 1/800 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

Thus

$$A(n, p) < \frac{n-1}{2(1-1/n)} - \tfrac{1}{2}(n+1) + \frac{1}{n} = -\tfrac{1}{2} + \frac{1}{n} < 1,$$

$$A(n, p) > \tfrac{1}{2}(n-1) - \tfrac{1}{2}(n+1) + \frac{1}{n} = -1 + \frac{1}{n} > -1. \qquad \square$$

**Proof of Theorem 3.2.** Let $e = np - \lfloor np \rfloor$. Then by Lemma 3.1 and (3.2),

$$\begin{aligned} d(n, p) &= np + \frac{np}{n} d(np, p) + \frac{n-np}{n} d(n-np, p) \\ &= (\lfloor np \rfloor + e) + \frac{\lfloor np \rfloor + e}{n}\left(d(\lfloor np \rfloor, p) + \frac{1}{2}\frac{e}{1-p}\right) \\ &\quad + \frac{n - \lfloor np \rfloor - e}{n}\left(d(n - \lfloor np \rfloor, p) - \tfrac{1}{2}\frac{e}{1-p}\right). \end{aligned}$$

With (3.1) we obtain

$$A(n, p) = R + \frac{\lfloor np \rfloor}{n} A(\lfloor np \rfloor, p) + \frac{n - \lfloor np \rfloor}{n} A(n - \lfloor np \rfloor, p)$$

where

$$R = e + \frac{e}{n}\frac{\lfloor np \rfloor + np - 1}{2(1-p)} - \frac{e}{n}\frac{2n - 1 - np - \lfloor np \rfloor}{2(1-p)} = \frac{e\lfloor np \rfloor}{n(1-p)}.$$

Therefore,

$$|A(n, p)| < \frac{\lfloor np \rfloor}{n} |A(\lfloor np \rfloor, p)| + \frac{n - \lfloor np \rfloor}{n} |A(n - \lfloor np \rfloor, p)| + \frac{p}{1-p}. \tag{A.1}$$

We will now prove the theorem by induction. For $n < 1/p$, by Lemma A.2, $|A(n, p)| < 1 \leqslant 2 \log n$. So assume $n \geqslant 1/p$ and the theorem holds for all $k < n$. Applying this hypothesis to (A.1) gives

$$|A(n, p)| < 2 \max\{\log\lfloor np \rfloor, \log(n - \lfloor np \rfloor)\} + \frac{p}{1-p} < 2(\log(n - \lfloor np \rfloor) + p) < 2 \log n,$$

where the last inequality follows from Lemma A.1. $\square$

Let

$$\alpha(y, p) = s(y, p) - s(\lfloor n \rfloor, p), \qquad \text{and} \qquad \beta(y, p) = s(y, p) - s(\lceil y \rceil, p).$$

**Lemma A.3.** *For* $y \geqslant 1$,

$$0 \leqslant \alpha(y, p) < \frac{1}{\lfloor y \rfloor H(p)},$$

*and*

$$-\frac{1}{yH(p)} < \beta(y, p) \leqslant 0.$$

**Proof.**

$$\alpha(y, p) = \frac{\log y - \log\lfloor y \rfloor}{H(p)} = \frac{1}{H(p)} \log\left(1 + \frac{y - \lfloor y \rfloor}{\lfloor y \rfloor}\right) < \frac{1}{H(p)\lfloor y \rfloor}$$

(since $\log(1+z) \leqslant z \ \ \forall z > -1$). On the other hand $s(y, p)$ is monotone increasing in $y$ so that nonnegativity of $\alpha(y, p)$ follows. The second inequality of the lemma is proved in a similar way:

$$\beta(y, p) = \frac{\log y - \log\lceil y\rceil}{H(p)} = \frac{1}{H(p)}\log\left(1 - \frac{\lceil y\rceil - y}{\lceil y\rceil}\right) > -\frac{1}{yH(p)}. \qquad \square$$

Let $B(n, p) = s(n, p) - S(n, p)$.

**Proof of Theorem 3.6.** Let $e = np - \lfloor np\rfloor$. By Lemma 3.5,

$$\begin{aligned} s(n, p) &= 1 + \frac{\lfloor np\rfloor + e}{n}[s(\lfloor np\rfloor, p) + \alpha(np, p)] \\ &\quad + \frac{n - \lfloor np\rfloor - e}{n}[s(n - \lfloor np\rfloor, p) + \beta(n - np, p)] \\ &= 1 + \frac{\lfloor np\rfloor}{n}s(\lfloor np\rfloor, p)\frac{n - \lfloor np\rfloor}{n}s(n - \lfloor np\rfloor, p) \\ &\quad + \frac{\lfloor np\rfloor}{n}\alpha(np, p) + \frac{n - \lfloor np\rfloor}{n}\beta(n - np, p) + R \end{aligned} \tag{A.2}$$

where

$$R = \frac{e}{n}[s(np, p) - s(n - np, p)] = \frac{e}{nH(p)}\log\frac{p}{1-p}. \tag{A.3}$$

With (3.3) we obtain

$$\begin{aligned} B(n, p) &= \frac{\lfloor np\rfloor}{n}B(\lfloor np\rfloor, p) + \frac{n - \lfloor np\rfloor}{n}B(n - \lfloor np\rfloor, p) + \frac{\lfloor np\rfloor}{n}\alpha(np, p) \\ &\quad + \frac{n - \lfloor np\rfloor}{n}\beta(n - np, p) + R. \end{aligned} \tag{A.4}$$

(i) We will now show that $B(n, p) < 1/(pH(p)) \log(4/p)$ by inductively showing that

$$B(n, p) < \frac{n-1}{npH(p)}. \tag{A.5}$$

For $n < 1/p$, $B(n, p) = (\log n)/H(p) - \frac{1}{2}(n+1) + 1/n$. Since $H > 0$, (A.5) will be proved if we show that

$$\log n + \frac{H(p)}{n} < \frac{n-1}{np}.$$

Since $p < 1/n$, the right-hand side is greater than $n - 1$. As for the left-hand side, Remark 3.4 implies

$$\begin{aligned} H(p) < H\left(\frac{1}{n}\right) &= -\frac{1}{n}\log\frac{1}{n} - \frac{n-1}{n}\log\frac{n-1}{n} \\ &= \log n - \frac{n-1}{n}\log(n-1) < \log n. \end{aligned}$$

This is suffices to show that $(1 + 1/n) \log n < n - 1$, and this holds for $n \geqslant 4$ since $((n+1)/(n-1)) \log n < 2 \log n \leqslant n$.

Suppose now that $n \geqslant 1/p$ and that (A.5) holds for all $k < n$. Noting that $R \leqslant 0$ and using Lemma A.3 with $y = np \geqslant 1$, we obtain from (A.4)

$$B(n, p) < \frac{\lfloor np\rfloor}{n}B(\lfloor np\rfloor, p) + \frac{n - \lfloor np\rfloor}{n}B(n - \lfloor np\rfloor, p) + \frac{1}{nH(p)}. \tag{A.6}$$

Substituting (A.5) in (A.6) we obtain

$$B(n, p) < \frac{\lfloor np \rfloor - 1}{npH(p)} + \frac{n - \lfloor np \rfloor - 1}{npH(p)} + \frac{1}{nH(p)} = \frac{1}{nH(p)}\left(\frac{n-2}{p} + 1\right) < \frac{n-1}{npH(p)}.$$

(ii) We will now show that $B(n, p) \geqslant (-1/(pH(p))\log(4/p)$ by inductively showing that

$$B(n, p) > -\frac{n-1}{npH(p)}\log\frac{4(1-p)}{p}. \tag{A.7}$$

Noting in (A.3) that $R > -\log((1-p)/p)/(nH(p))$ and using Lemma A.3, we obtain from (A.4):

$$B(n, p) > \frac{\lfloor np \rfloor}{n}B(\lfloor np \rfloor, p) + \frac{n - \lfloor np \rfloor}{n}B(n - \lfloor np \rfloor, p) - \frac{n - \lfloor np \rfloor}{n}\frac{1}{(n-np)H(p)}$$
$$- \frac{1}{nH(p)}\log\frac{1-p}{p}$$
$$\geqslant \frac{\lfloor np \rfloor}{n}B(\lfloor np \rfloor, p) + \frac{n - \lfloor np \rfloor}{n}B(n - \lfloor np \rfloor, p) - \frac{1}{nH(p)}\log\frac{4(1-p)}{p}. \tag{A.8}$$

For $n < 1/p$, $B(n, p) = (\log n)/H(p) - \frac{1}{2}(n+1) + 1/n$. Since $H > 0$, (A.7) will be proved if we show that by Remark 3.4, $H(p) < H(1/n)$, and therefore it suffices to show that

$$\log n - \frac{n+1}{n}\left(-\frac{1}{n}\log\frac{1}{n} - \frac{n-1}{n}\log\frac{n-1}{n}\right) > -(n-1)\log\frac{4(1-p)}{p}.$$

The inequality follows since the left-hand side is equal to

$$\log n - \tfrac{1}{2}(n+1)\left(\log n - \frac{n-1}{n}\log(n-1)\right)$$
$$= \tfrac{1}{2}(n-1)\left(\left(1 + \frac{1}{n}\right)\log(n-1) - \log n\right)$$
$$> \tfrac{1}{2}(n-1)(\log(n-1) - \log n) > -\tfrac{1}{2} > -(n-1)\log\frac{4(1-p)}{p}.$$

Suppose now that $n \geqslant 1/p$ and that (A.7) holds for all $k < n$. Substituting (A.7) in (A.8), we obtain

$$B(n, p) > -\frac{\lfloor np \rfloor - 1}{npH(p)}\log\frac{4(1-p)}{p} - \frac{n - \lfloor np \rfloor - 1}{npH(p)}\log\frac{4(1-p)}{p} - \frac{1}{nH(p)}\log\frac{4(1-p)}{p}$$
$$= \frac{1}{nH(p)}\log\frac{4(1-p)}{p}\left(-\frac{n-2}{p} - 1\right) > -\frac{n-1}{npH(p)}\log\frac{4(1-p)}{p}. \qquad \square$$

## References


Cameron, S.H., and Narayanamurthy, S.G. (1964), "A search problem", *Operations Research* 12, 623–629.

Fredman, M.L., and Knuth, D.E. (1974), "Recurrence relations based on minimization", *Journal of Mathematical Analysis and Applications* 48, 534–559.

Gilbert, E.N., and Moore, E.F. (1959), "Variable length binary encodings", *The Bell System Technical Journal* 38, 933–968.

Hassin, R. (1984), "A dichotomous search for a geometric random variable", *Operations Research* 32, 423–439.

Hu, T.C., and Tucker, A.C. (1971), "Optimal computer-search trees and variable-length alphabetical codes", *SIAM Journal of Applied Mathematics* 21, 514–532.

Hu, T.C., and Wachs, M.L. (1987), "Binary search on a tape", *SIAM Journal on Computing* 16, 573–590.

Kikuta, K. (1990), "A 1-dimensional search with travel cost", *Journal of the Operational Research Society of Japan* 33, 262.

Mehlhorn, K. (1984), *Data Structures and Algorithms: Sorting and Searching*, Vol. 1, Springer-Verlag, Berlin.

Morris, R. (1969), "Some theorems on sorting", *SIAM Journal of Applied Mathematics* 17, 1–6.

Murakami, S. (1971), "A dichotomous search", *Journal of the Operational Research Society of Japan* 14, 127–142.

Murakami, S. (1976), "A dichotomous search with travel cost", *Journal of the Operational Research Society of Japan* 19, 245–254.