

L. R. Johnson
M. H. McAndrew

On Ordered Retrieval from an Associative Memory

In a recent issue of this Journal, Seeber and Lindquist¹ describe an algorithm for ordered retrieval of the items stored in an associative memory. The basic method employs one or more memory interrogations per item retrieved. In discussing the average number of interrogations per item required for the case of N random binary numbers, they employ a formulation that is difficult to evaluate for large numbers and large N . The main purpose of this Communication is to extend their analysis for the case when N is large; toward this end, it is instructive and interesting in itself to generalize the algorithm for an arbitrary radix r , where $r \geq 2$. The presentation is in two parts. The first reviews the algorithm, exhibits a small example for numbers in a radix 4 representation, outlines the analytical approach, and comments on the nature of the results. The second presents a mathematical justification of the results.

General discussion

Let $u(r, N)$ denote the expected interrogations per retrieval in the ordered retrieval of N random numbers, where the numbers are represented in radix r . It will be assumed that each of the N numbers to be retrieved contains infinitely many digits. For any fixed N , therefore, the numbers may be considered distinct, i.e., as having been sampled without replacement. This simplifying assumption does not affect the practical significance of the results.

To illustrate the algorithm for radix 4, let the six numbers 023, 110, 122, 131, 302 and 320 be stored, in any order, as rows in a 6×3 matrix. During the course of ordered retrieval, the matrix is interrogated on arguments of the general form $m m m$, $d m m$, $d d m$, or $d d d$, where d denotes a radix 4 digit and m denotes a mask symbol. By convention, all rows match the argument $m m m$. For other arguments, an interrogation leads to a ternary outcome, that is, we have *no* matching row, *one* matching row, or *multiple* matching rows. After each interrogation yielding multiple matches, the leftmost m symbol is replaced by a 0. After every interrogation yielding no matches or one match, a 1 is added to the rightmost digit of the argument and the radix 4 carry is propa-

gated (if any exists); all 0's generated by the carry operation are replaced by m 's. At each interrogation with one match, the matching number is retrieved from the matrix. The algorithm starts with the argument $m m m$ and terminates when carry overflow reproduces this argument. The complete procedure for the six given numbers is shown in Table 1. Thirteen interrogations are required; six of these result in retrievals, as may be seen in the last column of the table.

It is suggestive to depict each of the thirteen required interrogations by a tree node, as in Figure 1. Each node represents the outcome of one interrogation and is identified by the argument employed in the interrogation. The leaves in the tree are associated with 0 or 1 match conditions and other nodes with multiple matches. The number shown in parentheses at each leaf with match condition 1 is the number retrieved by that interrogation.

Table 1 Procedure for ordered retrieval of six numbers for example of a radix 4 representation.

Interrogation number	Argument	Match condition	Number retrieved
1	$m m m$	>1	none
2	$0 m m$	1	023
3	$1 m m$	>1	none
4	$1 0 m$	0	none
5	$1 1 m$	1	110
6	$1 2 m$	1	122
7	$1 3 m$	1	131
8	$2 m m$	0	none
9	$3 m m$	>1	none
10	$3 0 m$	1	302
11	$3 1 m$	0	none
12	$3 2 m$	1	320
13	$3 3 m$	0	none
	$m m m$		

In the radix 4 case, each multiple match on level i of the tree implies four interrogations on the level $i + 1$. In a radix r generalization, each multiple match will likewise imply r more interrogations. If S denotes the number of required interrogations and M the number of multiple matches in a radix r example of N items, it is clear from the nature of the tree that

$$S = 1 + rM = 1 + r \sum_{i=0}^{\infty} M_i. \quad (1)$$

Thus the problem reduces to one of estimating the number of multiple matches, M_i , at tree level i .

Let N numbers be drawn at random and with equal probability from the series $0, 1, 2, \dots, n-2, n-1$. The probability that a given integer will occur in the sample with frequency k is given by the term $B(k; N, 1/n)$ of a binomial distribution.² By subtraction, the probability $B(k > 1; N, 1/n)$ of event $k > 1$ is equal to $1 - B(0; N, 1/n) - B(1; N, 1/n)$. As the given integer can be chosen in n ways, multiples are expected in $nB(k > 1; N, 1/n)$ cases.

At the i th level of the retrieval tree, n takes on the value r^i , where i ranges from zero to infinity. Thus if $m(r, N)$ denotes the expectation of M , it is seen that

$$m(r, N) = \sum_{i=0}^{\infty} r^i \{ 1 - B(0; N, 1/r^i) - B(1; N, 1/r^i) \}.$$

In explicit computational form this becomes

$$m(r, N) = \sum_{i=0}^{\infty} r^i \{ 1 - (1 - r^{-i})^N - Nr^{-i}(1 - r^{-i})^{N-1} \}. \quad (2)$$

The expected number of interrogations for N numbers being $1 + rm(r, N)$, it follows that

$$u(r, N) = \frac{1}{N} + \frac{r}{N} m(r, N). \quad (3)$$

Although Eq. (2) is difficult to compute for all but small values of N , the Poisson approximation³ was found to yield very satisfactory results. There is little point in tabulating the function, however, because an approximation with error term of order $(\log N)/N$ is available in the limiting values of $u(r, N)$. Let $F(r, \alpha)$ denote the limit of $u(r, N)$ as N goes to infinity through values $r^{j+\alpha}$ for integral j . Then, as shown in Theorem 1 below,

$$u(r, N) = F(r, \alpha) + O(\log N/N). \quad (4)$$

Moreover, by Theorem 2, $F(r, \alpha)$ is periodic in α , is of period 1, and its mean over a cycle is $r/\log r$; that is,

$$\int_0^1 F(r, \alpha) d\alpha = r/\log r. \quad (5)$$

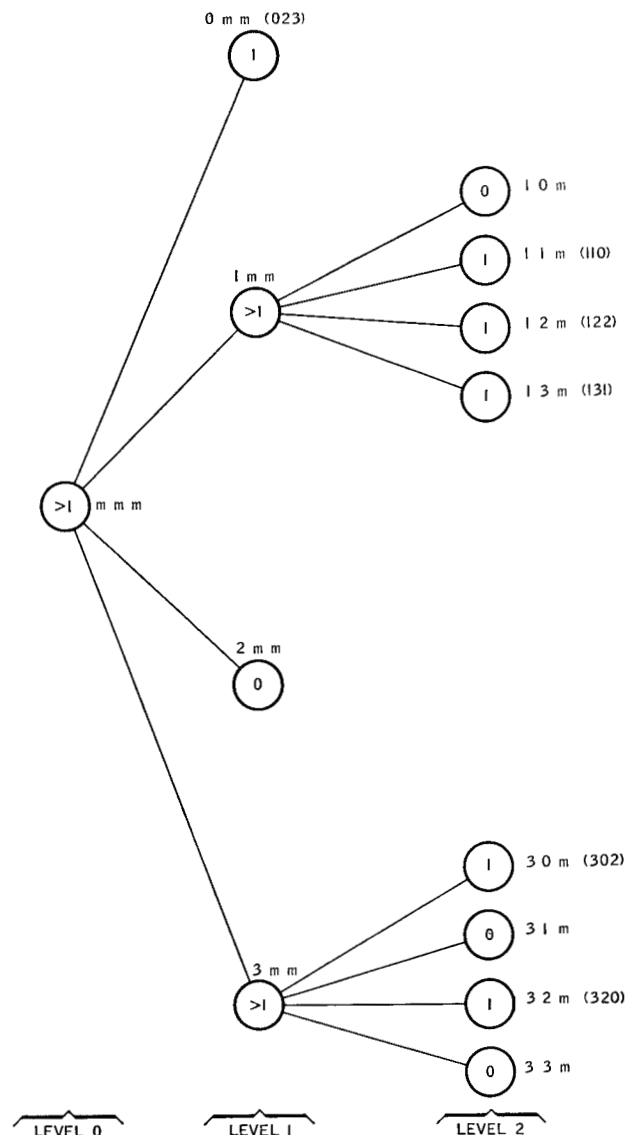


Figure 1 Retrieval tree for example of Table 1. The nodes encircle match conditions and are labelled to show relevant arguments. Retrieval numbers are enclosed in parentheses.

In Lemma 6 below an explicit form for $F(r, \alpha)$ is obtained, namely:

$$F(r, \alpha) = r \sum_{j=-\infty}^{\infty} g(r^{\alpha-j}),$$

where $g(\theta) = (1 - e^{-\theta})/\theta - e^{-\theta}$. From this, $F(r, \alpha)$ was computed for $r = 2, 3, 4, 8, 10, 16, 32, 64$ and for α ranging from 0 to 1 in intervals of 0.05. The results of the computation are displayed in Table 2. The column for $r = 4$, say, approximates $u(4, N)$ in a period from $N = r^j$ to

Table 2 Selected values of $F(r, \alpha)$, computed for $r = 2, 3, 4, 8, 10, 16, 32$ and 64 , and for values of α ranging from 0 to 1 in intervals of 0.05.

α	$r = 2$	3	4	8	10	16	32	64
0.05	2.88539	2.73141	2.88241	3.80851	4.32698	5.95426	10.44045	19.56421
0.10	2.88539	2.73134	2.88117	3.83588	4.37913	6.08546	10.78590	20.32025
0.15	2.88539	2.73121	2.88035	3.86422	4.42774	6.18873	10.98707	20.56146
0.20	2.88539	2.73102	2.88002	3.89082	4.46836	6.25349	11.01155	20.20603
0.25	2.88539	2.73081	2.88022	3.91319	4.49716	6.27207	10.84505	19.26207
0.30	2.88539	2.73059	2.88092	3.92922	4.51132	6.24090	10.49769	17.84543
0.35	2.88539	2.73038	2.88206	3.93739	4.50928	6.16148	10.00632	16.16481
0.40	2.88539	2.73020	2.88353	3.93687	4.49095	6.04081	9.43068	14.46993
0.45	2.88539	2.73008	2.88518	3.92764	4.45786	5.89107	8.84324	12.98299
0.50	2.88538	2.73002	2.88685	3.91049	4.41305	5.72832	8.31488	11.84801
0.55	2.88538	2.73002	2.88837	3.88702	4.36092	5.57043	7.90163	11.12280
0.60	2.88538	2.73009	2.88961	3.85949	4.30674	5.43462	7.63721	10.80629
0.65	2.88538	2.73022	2.89043	3.83061	4.25608	5.33495	7.53329	10.87233
0.70	2.88538	2.73041	2.89075	3.80329	4.21421	5.28062	7.58520	11.28894
0.75	2.88539	2.73062	2.89055	3.78030	4.18542	5.27532	7.77849	12.02183
0.80	2.88539	2.73084	2.88985	3.76398	4.17253	5.31752	8.09302	13.03050
0.85	2.88539	2.73105	2.88871	3.75597	4.17662	5.40153	8.50419	14.26155
0.90	2.88539	2.73122	2.88725	3.75703	4.19698	5.51858	8.98249	15.64191
0.95	2.88539	2.73135	2.88560	3.76699	4.23133	5.65791	9.49267	17.07325
1.00	2.88539	2.73142	2.88393	3.78475	4.27612	5.80741	9.99379	18.43003
average	2.88539	2.73071	2.88539	3.84718	4.34294	5.77077	9.23324	15.38873

$N = r^{j+1}$, (from, for example, 16 to 64, 64 to 256, etc.). It is evident that $u(2, N)$ varies by a very small amount but that periodic variation increases with r . From the standpoint of retrieval, radices 2 and 4 may be considered equal in efficiency but 8 is considerably less efficient. Among integral radices, 3 proves the most efficient choice of all.

Mathematical justification: theorems

• Lemma 1

For $N > 0$ and $0 \leq x \leq 0.1$,

$$-0.56Nx^2 \leq (1-x)^N - e^{-Nx} \leq 0.$$

• Proof

$N \log(1-x) = -Nx - Nx^2/2 - Nx^3/3 - \dots$. Hence an upper bound for $N \log(1-x)$ is $-Nx$; for a

lower bound we have

$$\begin{aligned} N \log(1-x) &\geq -Nx - Nx^2/2 - Nx^3/2 - \dots \\ &= -Nx - Nx^2/2(1-x) \\ &\geq -Nx - 0.56Nx^2, \text{ since } x \leq 0.1. \end{aligned}$$

Taking exponentials we have

$$\begin{aligned} \exp(-Nx) &\geq (1-x)^N \geq \exp(-Nx - 0.56Nx^2) \\ &\geq (1 - 0.56Nx^2) \exp(-Nx) \\ &\geq \exp(-Nx) - 0.56Nx^2. \end{aligned}$$

• Lemma 2

For $0 \leq x \leq 1$ and $N \geq 2$,

$$\begin{aligned} \text{(i)} \quad & |(1-x)^N - (1-x)^{N-1}| \leq 1/N, \\ \text{(ii)} \quad & |(1-x)^N - (1-x)^{N-1}| \leq x. \end{aligned}$$

• *Proof*

(i) Let $f(x) = (1 - x)^N - (1 - x)^{N-1}$. Then $f(0) = f(1) = 0$ and, for $0 < x < 1$, $f'(x) = 0$ only at $x = 1/N$. Hence

$$\begin{aligned} |f(x)| &\leq |f(1/N)| \\ &= (1 - 1/N)^{N-1}/N \\ &< 1/N. \end{aligned}$$

(ii) The result follows since

$$(1 - x)^N - (1 - x)^{N-1} = -x(1 - x)^{N-1}.$$

• *Lemma 3*

For $0 \leq x \leq 1$, and $N \geq 2$,

$$|(1 - x)^N - e^{-Nx}| \leq 1/N.$$

• *Proof*

Let $f(x) = (1 - x)^N - e^{-Nx}$, then $f(0) = 0$ and $f(1) = -e^{-N}$. At any turning point x_0 of $f(x)$, in $0 \leq x \leq 1$, $-N(1 - x_0)^{N-1} + Ne^{-Nx_0} = 0$, i. e., $e^{-Nx_0} = (1 - x_0)^{N-1}$. Hence $f(x_0) = (1 - x_0)^N - (1 - x_0)^{N-1}$ and so, by Lemma 2(i), $|f(x_0)| \leq 1/N$. Hence, for $0 \leq x \leq 1$, $|f(x)| \leq \max(1/N, e^{-N}) \leq 1/N$.

• *Lemma 4*

Let $u(r, N)$ be defined as in the general discussion. Then for $N \geq 10$, $r > 1$,

$$\begin{aligned} u(r, N) &= \frac{1}{N} + \frac{r}{N} \sum_{i=0}^{\infty} r^i \left(1 - e^{-N/r^i} \right. \\ &\quad \left. - \frac{N}{r^i} e^{-N/r^i} \right) + E_1, \end{aligned}$$

where

$$|E_1| < \frac{3r(\log_r N + 1)}{N} + \frac{2.12r^2}{N(r - 1)}.$$

• *Proof*

$$\begin{aligned} NE_1/r &= \sum_{i=0}^{\infty} r^i \{ 1 - (1 - 1/r^i)^N - N(1/r^i) \\ &\quad \cdot (1 - 1/r^i)^{N-1} \} - \sum_{i=0}^{\infty} r^i \left(1 - e^{-N/r^i} - \frac{N}{r^i} e^{-N/r^i} \right) \\ &= - \sum_{i=0}^{\infty} (r^i + N) \{ (1 - 1/r^i)^N - e^{-N/r^i} \} \\ &\quad + \sum_{i=0}^{\infty} N \{ (1 - 1/r^i)^N - (1 - 1/r^i)^{N-1} \}. \end{aligned}$$

Using the estimates of Lemma 2(i) and Lemma 3 for terms

of these series with $r^i < N$, and the estimates of Lemma 1 and Lemma 3(ii) for terms with $r^i \geq N$, we obtain

$$\begin{aligned} |NE_1/r| &\leq \left\{ \sum_{r^i < N} 2N(1/N) + \sum_{r^i \geq N} 2r^i(0.56N/r^{2i}) \right. \\ &\quad \left. + \sum_{r^i < N} N(1/N) + \sum_{r^i \geq N} N(1/r^i) \right\} \\ &= \sum_{0 \leq i < \log_r N} 3 + \sum_{r^i \geq N} 2.12N/r^i \\ &< 3(\log_r N + 1) + \frac{2.12r}{r - 1}. \end{aligned}$$

• *Lemma 5*

Let $g(\theta) = (1 - e^{-\theta})/\theta - e^{-\theta}$. Then for $\theta > 0$, $0 < g(\theta) < 1/\theta$.

• *Proof*

Let $f(\theta) = \theta g(\theta) = 1 - e^{-\theta} - \theta e^{-\theta}$. Then $f'(\theta) = \theta e^{-\theta} > 0$. Hence, for $\theta > 0$, $0 = f(0) < f(\theta) < f(\infty) = 1$.

• *Lemma 6*

Let $N = r^{k+\alpha}$ where k is integral, $r > 1$, and $0 \leq \alpha \leq 1$. Then

$$f(r, N) = F(r, \alpha) + E_2,$$

where

$$F(r, \alpha) = r \sum_{i=-\infty}^{\infty} g(r^{\alpha-i}),$$

$$f(r, N) = \frac{r}{N} \sum_{i=0}^{\infty} r^i \left(1 - e^{-N/r^i} - \frac{N}{r^i} e^{-N/r^i} \right),$$

$$g(\theta) = \frac{1 - e^{-\theta}}{\theta} - e^{-\theta}, \text{ and}$$

$$|E_2| < \frac{r}{(r - 1)N}.$$

• *Proof*

$$\begin{aligned} f(r, N) &= \frac{r}{N} \sum_{i=0}^{\infty} r^i \left(1 - e^{-N/r^i} - \frac{N}{r^i} e^{-N/r^i} \right) \\ &= r \sum_{i=0}^{\infty} g(N/r^i) = r \sum_{i=0}^{\infty} g(r^{k+\alpha-i}) = r \sum_{j=-k}^{\infty} g(r^{\alpha-j}). \end{aligned}$$

Hence,

$$E_2 = f(r, N) - F(r, \alpha) = -r \sum_{j=-\infty}^{-k-1} g(r^{\alpha-j}).$$

By Lemma 5,

$$|E_2| < r \sum_{j=-\infty}^{-k-1} r^{j-\alpha} = r/N(r - 1).$$

• **Theorem 1**

Let $N = r^{k+\theta}$ for integral k and $0 \leq \alpha < 1$. Let $F(r, \alpha)$ be defined as in Lemma 6. Then for $N \geq 10$ and $r > 1$,

$$|u(r, N) - F(r, \alpha)|$$

$$< \frac{3r(\log_r N + 1)}{N} + \frac{2.12r^2 + r}{N(r-1)} + \frac{1}{N}.$$

• **Proof**

The result follows directly from Lemmas 4 and 6.

• **Theorem 2**

$F(r, \alpha)$ is a periodic function of α of period 1; further

$$\int_0^1 F(r, \alpha) d\alpha = \frac{r}{\log r}.$$

• **Proof**

The periodicity of $F(r, \alpha)$ is clear from its expression as $r \sum_{i=-\infty}^{\infty} g(r^{\alpha-i})$. The value of the integral follows

directly from term-by-term integration of this series since

$$\int_0^1 g(r^{\alpha-i}) d\alpha = \int_{r^{-i}}^{r^{1-i}} \frac{g(x)}{x \log r} dx$$

and

$$\int_0^{\infty} \frac{g(x)}{x} dx = 1.$$

Acknowledgment

The authors are grateful to R. R. Seeber and A. B. Lindquist for their interest and for their helpful comments on the work in manuscript.

References

1. R. R. Seeber and A. B. Lindquist, "Associative Memory with Ordered Retrieval," *IBM Journal* **6**, 126-136 (1962).
2. See, for example, W. Feller, *Probability Theory and Its Applications*, Vol. I, 2nd ed., John Wiley & Sons, New York, 1957, p. 137.
3. Feller, op. cit., pp. 142-145.

Received December 26, 1963