A Combinatorial Problem Related to Interleaved Memory Systems

G. J. BURNETT AND E. G. COFFMAN, JR.

The Pennsylvania State University, University Park, Pennsylvania

ABSTRACT. A combinatorial problem arising from the analysis of a model of interleaved memory systems is studied. The performance measure whose calculation defines this problem is based on the distribution of the number of modules in operation during a memory cycle, assuming saturated demand and an arbitrary but fixed number of modules.

In general terms the problem is as follows. Suppose we have a Markov chain of n states numbered $0, 1, \cdots, n - 1$. For each i assume that the one-step transition probability from state i to state $(i + 1) \bmod n$ is given by the parameter α and from state i to any other state is $\beta = (1 - \alpha)/(n - 1)$. Given an initial state, the problem is to find the expected number of states through which the system passes before returning to a state previously entered. The principal result of the paper is a recursive procedure for computing this expected number of states. The complexity of the procedure is seen to be small enough to enable practical numerical studies of interleaved memory systems.

KEYWORDS AND PHRASES: interleaved memory systems, modular memory systems, memory performance analysis, storage systems, memory accessing, combinatorics

CR CATEGORIES: 5.39, 6.34

1. *Introduction*

In a recent paper [1] a model of interleaved memory systems was devised and analyzed. In the analysis, separate models of instruction and data addressing were proposed in order to consider possible improvements in the average memory bandwidth, defined to be the average number of memory modules in operation per memory cycle. The numerical studies demonstrated the trade-off between instruction and memory cycle speeds and also showed that significant increases in the average memory bandwidth can be obtained by separately grouping instruction and data requests when accessing the memory.

These numerical studies were based on mathematical analyses of instruction and data addressing models. The analysis of data addressing in an interleaved memory system is of particular interest as a combinatorial problem, and it forms the subject of this paper.

This research was supported in part by NSF Grant GJ-28290.

Author's address: The Pennsylvania State University, Computer Science Department, University Park, PA 16802.

2. *The Model*

The model to be analyzed is shown in Figure 1. There are n identical modules, each capable of reading or writing one word per memory cycle. We shall assume that the modules operate synchronously and with identical cycle times. In practice, the request sequence $r_1, \cdots, r_k, \cdots$ in the request queue contains conventional storage addresses; however, for our purposes only the module number from the address is of interest. Thus we will consider the requests r_i, $i = 1, 2, \cdots$, to be integers from the set $S_n = \{0, 1, \cdots, n-1\}$. The scanner operates by admitting new requests to service until it attempts to assign a request to a busy memory module. To do this, prior to the start of a given memory cycle (i.e. during the previous memory cycle) the scanner inspects the request queue beginning with the request rejected for the previous cycle, and determines the maximum length sequence of distinct module requests. That is, it scans the queue to the first repetition of a module request. The memory requests in this maximum length sequence are then sent to the appropriate memory modules so that they will be active in the next memory cycle. During this next memory cycle the above process is repeated to obtain the requests served in the subsequent cycle.

We shall assume that the request queue always contains more than n requests when inspected. In effect, the queue will always be saturated and our interest will be in characterizing system capacity. Clearly, the effectiveness of the system is determined by the probability mass function (pmf) governing the number of distinct requests served in each memory cycle. This pmf is to be found assuming that the probability model is the one defined below. (This model may be regarded as a generalization of a simple model studied by Hellerman [2].)

The request sequences are modeled by the parameter α whose meaning is given as follows. The first request in the request queue addresses a module at random. Thereafter, the ith request, $i \geq 2$, addresses the next module in sequence (modulo n) with stationary probability α, and addresses any one of the modules out of sequence with probability $\beta = (1-\alpha)/(n-1)$. Formally, let $r_1, \cdots, r_i, \cdots$ denote the contents of the request queue. Then

$$\begin{aligned} &\Pr(r_1 = m) = 1/n, \qquad m \in S_n = \{0, 1, \cdots, n-1\}; \\ &\Pr(r_{i+1} = (r_i + 1) \bmod n) = \alpha, \qquad i = 1, 2, \cdots; \\ &\Pr(r_{i+1} = m) = \beta, \qquad m \in S_n,\ m \neq (r_i + 1) \bmod n,\ i = 1, 2, \cdots. \end{aligned} \tag{1}$$

For example, in a system having $n \geq 7$ modules, the initial subsequence 0,5,6,2,3 would have probability $(1/n)\alpha^2\beta^2$.

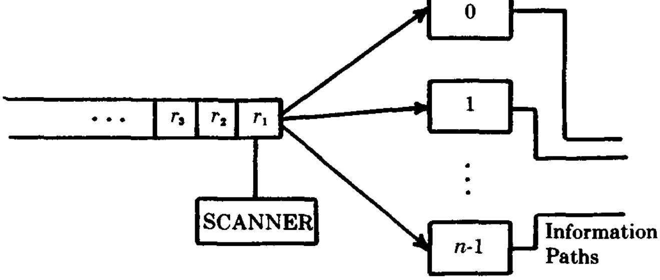

FIG. 1

Let w_i, $i = 1, 2, \cdots$, be the random variable whose value is the number of requests served on the ith memory cycle. A little reflection convinces us that the distribution for w_i is independent of the value of the first request inspected in the $(i-1)$th cycle. It follows that the distributions of the w_i's are the same for all i. Thus we shall drop the subscript i and adopt the notation $P_n(w \leq k)$, $1 \leq k \leq n$, as the common, cumulative distribution function for an n-module system. The mean value of this distribution is given by

$$B_n = \sum_{k=1}^{n} P_n(w \geq k) \tag{2}$$

with the units of words per memory cycle. The remainder of the paper is devoted to the computation of B_n, the average memory bandwidth.

In general terms the problem can be restated as follows. Suppose we have a Markov chain of n states, $0, 1, \cdots, n-1$. For each i suppose that the one-step transition probability from state i to state $(i+1) \bmod n$ is given by the parameter α and from state i to any other state is $\beta = (1-\alpha)/(n-1)$. Given an initial state, the problem is to find the expected number of states through which the system passes before returning to a state previously entered.

3. *Computation of B_n*

We begin by observing that $P_n(w \geq k)$ is simply the probability that the first k requests found in the request queue are distinct. Next, in computing B_n it is convenient to have the following characterization of request sequences. For the sequence, $\underline{r} = r_1, \cdots, r_k$, we shall say that (r_i, r_{i+1}), $1 \leq i < k$, is an *α-transition* if $r_{i+1} = (r_i + 1) \bmod n$; otherwise, it will be called a *β-transition.*

Assuming an n-module system, we let $c_n(j, k)$ denote the number of k-length sequences of distinct integers for which the number of α-transitions is j and which begin with the request $r_1 = 0$. We have chosen the first request to be 0 merely for convenience; clearly, the number of such sequences beginning with $r_1 = i$ is the same as the number beginning with any $j \neq i$. It follows easily that we can write

$$B_n = \sum_{k=1}^{n} \sum_{j=0}^{k-1} \alpha^j \beta^{k-j-1} c_n(j, k) \tag{3}$$

and our problem is now to find the numbers $c_n(j, k)$.

We first register the fact that there are $\binom{k-1}{j}$ possible orderings of the α and β-transitions in k-length sequences with j α-transitions. Let $c_n^0(j, k)$ denote the number of sequences of distinct integers for which the first j transitions are α-transitions and the remaining $k - j - 1$ are β-transitions. For example, if $j = 6$ and $k = 10 \leq n$, then $c_n^0(6, 10)$ denotes the number of sequences of length 10 having the form $0, 1, 2, 3, 4, 5, 6, r_8, r_9, r_{10}$ in which the only α-transitions are the 6 present in the initial subsequence $0, 1, 2, 3, 4, 5, 6$.

Note immediately that $c_n(0, k) = c_n^0(0, k)$ and $c_n(k-1, k) = c_n^0(k-1, k)$. Also, since there is only one sequence of k consecutive integers beginning with 0 we have

$$c_n^0(k-1, k) = 1, \qquad 1 \leq k \leq n. \tag{4}$$

The total number of k-length sequences of distinct integers drawn from S_n and beginning with 0 is $(n-1)_{k-1} = (n-1)(n-2)\cdots(n-k+1)$. (We define $(n-1)_0 = 1$.) Hence,

$$(n-1)_{k-1} = \sum_{j=0}^{k-1} c_n(j,k). \tag{5}$$

The following result significantly simplifies the computation of B_n. It follows from the fact (which we shall prove) that the number of k-length sequences with one ordering of j α-transitions and $k-j-1$ β-transitions is the same as that for any other ordering of the same numbers of α and β-transitions.

THEOREM 1.

$$c_n(j,k) = \binom{k-1}{j} c_n^{0}(j,k). \tag{6}$$

PROOF. Let E_q denote the set of k-length sequences of distinct integers which have the qth $\left(q = 1, 2, \cdots, \binom{k-1}{j}\right)$ ordering of j α-transitions and $k-j-1$ β-transitions.[1] Let q' index an ordering that differs from the qth only by a single transposition of adjacent elements (transitions). It is well known that any two permutations of j α's and $k-j-1$ β's can be obtained from one another by sequences of transpositions of adjacent elements. Thus the theorem will follow easily if we can show that $|E_q| = |E_{q'}|$. To show this we will produce a one-to-one mapping from E_q onto $E_{q'}$.

Let $r_1, \cdots, r_k$ be a sequence in E_q and suppose that the ordering of transitions in sequences of E_q differs from that of sequences in $E_{q'}$ only in the ith and $(i+1)$th transitions. In particular, if $r_1', \cdots, r_k'$ is in $E_{q'}$ let us suppose that $(r_i, r_{i+1}) = (r_{i+1}', r_{i+2}') = \alpha$ and $(r_{i+1}, r_{i+2}) = (r_i', r_{i+1}') = \beta$. We map $r_1, \cdots, r_k$ into a new sequence $r_1', \cdots, r_k'$ as follows.

Let $r_i = m$ and $r_{i+2} = m'$. Then for each j $(1 \le j \le k)$, r_j' is obtained from r according to the integer mapping $f: S_n \to S_n$:

$$\begin{array}{ll}
m' \to m' & (m+1) \bmod n \to (m'-1) \bmod n \\
(m'+1) \bmod n \to (m'+1) \bmod n & (m+2) \bmod n \to (m+1) \bmod n \\
\cdot \qquad \cdot & \cdot \\
\cdot & \cdot \\
\cdot \qquad \cdot & \cdot \\
(m-1) \bmod n \to (m-1) \bmod n & \\
m \to m & (m'-1) \bmod n \to (m'-2) \bmod n
\end{array} \tag{7}$$

Figure 2 shows two examples for $j = 4$, $k = 6$, and $n = 10$. It is immediately evident from (7) that $(r_i', r_{i+1}') = \beta$, $(r_{i+1}', r_{i+2}') = \alpha$, and the r_j' are distinct. We now verify that $(r_j', r_{j+1}') = (r_j, r_{j+1})$ for all $j < i$ and $j > i+1$, in order to show that the new sequence is indeed always in $E_{q'}$.

In the following we assume $j < i$ or $j > i+1$, and hence $r_j \ne m$, $r_j \ne m+1$, $r_{j+1} \ne m+1$, and $r_{j+1} \ne m'$. Clearly, if r_j and r_{j+1} are both unchanged or if 1 is subtracted (mod n) from both, then $(r_j, r_{j+1}) = (r_j', r_{j+1}')$. When they are not

[1] We do *not* restrict the first element of the sequences in E_q.

	r_1		r_2		r_3		r_4		r_5		r_6		r_7	
qth ordering:	0	α	1	β	6	α	7	α	8	β	4	α	5	
														$i = 4,\ m = 7,\ m' = 4$
(q')th ordering:	9	α	0	β	6	α	7	β	3	α	4	α	5	
qth ordering:	4	β	1	α	2	α	3	β	5	α	6	α	7	
														$i = 3,\ m = 2,\ m' = 5$
(q')th ordering:	3	β	1	α	2	β	4	α	5	α	6	α	7	

FIG. 2

affected the same way only two cases can arise:

(a) $r_j \in \{(m+2) \bmod n, \cdots, (m'-1) \bmod n\}$ and
$r_{j+1} \in \{(m'+1) \bmod n, \cdots, m\}$,
in which case $r_j' = (r_j - 1) \bmod n$ and $r_{j+1}' = r_{j+1}$; or

(b) $r_j \in \{m', \cdots, (m-1) \bmod n\}$ and
$r_{j+1} \in \{(m+2) \bmod n, \cdots, (m'-1) \bmod n\}$,
in which case $r_j' = r_j$ and $r_{j+1}' = (r_{j+1} - 1) \bmod n$.

In either case it is readily verified that $(r_j, r_{j+1}) = (r_j', r_{j+1}') = \beta$ for all valid choices of r_j and r_{j+1}. Hence, $r_1', \cdots, r_k'$ is a sequence in $E_{q'}$.

It is easily seen that (7) generates a mapping of sequences $f' : E_q \rightarrow E_{q'}$ that is both one-to-one and onto. Since the inverse mapping is one-to-one onto we have also taken care of the assumption $(r_i, r_{i+1}) = \beta$ and $(r_{i+1}, r_{i+2}) = \alpha$.

Since we are now free to choose the ordering of transitions, we choose the one in which all of the α-transitions appear first. Restricting to sequences beginning with 0, (6) follows from the observation that there are $\binom{k-1}{j}$ possible orderings of j α-transitions and $k - j - 1$ β-transitions. ▌

Using Theorem 1 we can develop the main result.

THEOREM 2.

$$B_n = \sum_{k=1}^{n} \sum_{j=0}^{k-1} \binom{k-1}{j} x_{k-j,n-j}\, \alpha^j \beta^{k-j-1} \tag{8}$$

where

$$x_{k,n} = (n-1)_{k-1} - \sum_{j=1}^{k-1} \binom{k-1}{j} x_{k-j,n-j} \tag{9}$$

in which the summation is defined to be 0 if $k = 1$.

PROOF. First, it is easily seen that

$$c_n^0(j, k) = c_{n-1}^0(j-1, k-1), \qquad 1 \leq j \leq k-1 \leq n-1. \tag{10}$$

For by dropping the last α-transition and considering S_{n-1} instead of S_n, we obtain the same number of $(k-1)$-length sequences for the parameters $j - 1$ and

$n - 1$ as we had k-length sequences for j and n. One boundary condition is provided by (4):

$$c_n^0(0, 1) = 1; \tag{11}$$

and another is provided by the observation that in a sequence of two requests beginning with 0 and having a single β-transition, we must have $r_2 \neq 0$ and $r_2 \neq 1$:

$$c_n^0(0, 2) = n - 2. \tag{12}$$

From (10) we obtain

$$c_n^0(j, k) = c_{n-j}^0(0, k - j). \tag{13}$$

Next, using (5) and (6) we get

$$c_n(0, k) = c_n^0(0, k) = (n - 1)_{k-1} - \sum_{j=1}^{k-1} \binom{k-1}{j} c_n^0(j, k)$$

whereupon substitution of (13) gives

$$c_n^0(0, k) = (n - 1)_{k-1} - \sum_{j=1}^{k-1} \binom{k-1}{j} c_{n-j}^0(0, k - j).$$

Introducing the simpler notation $x_{k,n} \equiv c_n^0(0, k)$ we get (9). Using (6) and (13) in (3) we obtain (8). ∎

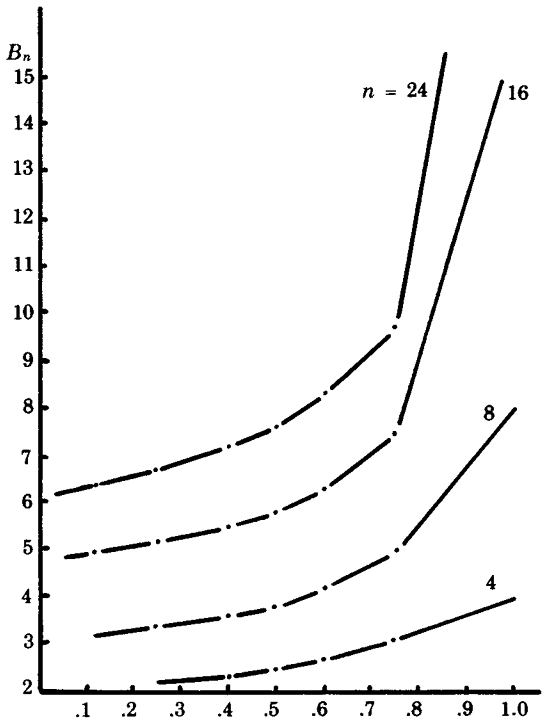

FIG. 3

4. *Final Remarks*

The complexity of a procedure for evaluating B_n for given α and n is dominated by the computation of the $x_{k,n}$. The $x_{k,n}$ are evaluated in the order $x_{11}, \cdots, x_{1n}$, $x_{22}, \cdots, x_{2n}, x_{33}, \cdots \cdots, x_{n-1,n}$, and require a total of $O(n^3)$ simple additions.

Figure 3 illustrates the behavior of B_n as a function of α for various n. The limiting case $\alpha = \beta = 1/n$ corresponds to Hellerman's random addressing model, while $\alpha = 1$ obviously gives $B_n = n$. For a more complete numerical study including different methods of handling instruction and data accessing, see [1] and [3].

ACKNOWLEDGMENT. The authors would like to express their thanks for the special efforts of Professor Ron Read in examining the manuscript and in providing valuable suggestions.

REFERENCES

1. BURNETT, G. J , AND COFFMAN, E G JR. A study of interleaved memory systems. Proc' AFIPS 1970 SJCC, Vol. 36, AFIPS Press, Montvale, N. J., pp 467–474.
2. HELLERMAN, H. *Digital Computer System Principles* McGraw-Hill, New York, 1967
3. BURNETT, G. J. Performance analysis of interleaved memory systems. Ph.D. Thesis, Elec. Eng. Dept., Princeton U., Princeton, N. J., Jan. 1970.

RECEIVED OCTOBER 1970; REVISED FEBRUARY 1972.