Expected Deadlock Time in a Multiprocessing System

KEVIN J. COMPTON AND CHINYA RAVISHANKAR

University of Michigan, Ann Arbor, Michigan

Abstract. We consider multiprocessing systems where processes make independent, Poisson distributed resource requests with mean arrival time 1. We assume that resources are not released. It is shown that the expected deadlock time is never less than 1, no matter how many processes and resources are in the system. Also, the expected number of processes blocked by deadlock time is one-half more than half the number of initially active processes. We obtain expressions for system statistics such as expected deadlock time, expected total processing time, and system efficiency, in terms of Abel sums. We derive asymptotic expressions for these statistics in the case of systems with many processes and the case of systems with a fixed number of processes. In the latter, generalizations of the Ramanujan Q-function arise. We use singularity analysis to obtain asymptotics of coefficients of generalized Q-functions.

Categories and Subject Descriptors: D.4.1 [**Operating Systems**]; Process Management—*deadlocks; multiprocessing/multiprogramming*; F.2.2 [**Analysis of Algorithms and Problem Complexity**]; Nonnumerical Algorithms and Problems—*computations on discrete structures*; G.2.1 [**Discrete Mathematics**]; Combinatorics—*generating functions*; G.2.2 [**Discrete Mathematics**]; Graph Theory—*path and circuit problems*

General Terms: Performance, Theory

Additional Key Words and Phrases: Asymptotic analysis, expected time analysis, singularity analysis

1. *Introduction*

Deadlock detection and resolution is a major issue in the design of multiprocessing systems (see Bic and Shaw [1988]). Although, in some systems (Unix, for example) deadlock is rare and the cost of resolution is low, in many others (such as database systems) the likelihood of deadlock may be quite high and resolution requires an expensive rollback and recovery. It would be useful to know under what circumstances deadlock is likely and (especially when resolution is costly) the expected time for the occurrence of deadlock. This paper presents a model of multiprocessing systems where processes make resource requests independently and with Poisson distributions of mean 1. We derive

K. J. Compton would like to express his appreciation to the Institut National de Recherche en Informatique et en Automatique and l'Ecole Polytechnique for partial support during the period this research was done. He would also like to thank the members of Projet Algo at INRIA for providing a supportive research environment and good camaraderie in the 1991–92 academic year.

Authors' present address: EECS Department, University of Michigan, Ann Arbor, MI 48109-2122.

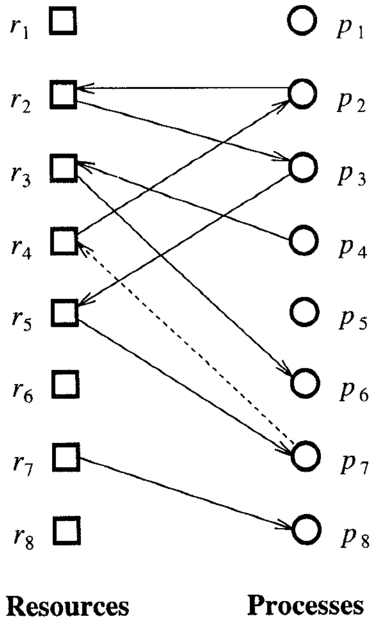

FIG. 1. A resource allocation graph.

exact and asymptotic expressions for system statistics such as expected time to deadlock, expected total processing time, and system efficiency. We make the simplifying assumption that resources are never released. Thus, our results may be viewed as upper bounds or bounds for an extreme case of system behavior.

Let us describe our model a little more precisely. A *multiprocessing system* is composed of two types of entities: processes and resources. *Processes* are the active entities of the system. They can change the system state by requesting new resources or releasing resources allocated to them. Resources are *serially reusable*: they may be reallocated once they are released. Examples of such resources are hardware units such as memory pages or printers, and software resources such as database locks. We do not examine systems with *consumable resources* such as messages, signals, and input data. We also assume that each process requests only one resource at a time.

A system state is represented by a *resource allocation graph*. This is a directed graph whose vertices are the processes and resources in the system. The graph is bipartite; edges are directed from resources to processes or processes to resources (see Figure 1).

A resource is *free* if it has not been allocated to a process. If a process p requests a free resource r, an edge is inserted in the resource allocation graph from r to p to indicate that r has been allocated to p and is no longer free. When p releases r the edge is erased. If p requests a resource r that is not free, an edge is drawn from p to r, indicating that p is waiting for r to be released. In this case, p becomes *inactive* or *blocked* and can make no more requests until r is released. Thus, an active process has out-degree 0 and an inactive process has out-degree 1. Deadlock occurs when a directed cycle appears in the resource allocation graph. Since all the processes on the directed cycle are blocked, the resources on the cycle can never be released. They become useless until the deadlock is detected and resolved. For example, if process p_7 requests resource r_4 in Figure 1, deadlock occurs because a cycle

$p_7 \to r_4 \to p_2 \to r_2 \to p_3 \to r_5 \to p_7$ results when the edge $p_7 \to r_4$ (indicated by the dotted arrow) is added. A more detailed description of the model is given in the following section.

Lynch [1994], in enumerating the number of cycle-free resource allocation graphs for given numbers of resources and processes, used the following necessary and sufficient conditions for a bipartite directed graph to be a resource allocation graph.

(i) Every vertex has out-degree at most one.
(ii) Every resource with non-zero in-degree has out-degree exactly one.

To see this, observe that a resource acquires an out-edge the first time it is requested; thereafter, it acquires only in-edges. A process, on the other hand, acquires in-edges until it acquires its first out-edge; thereafter, it is inactive. (For more general models of deadlock and associated references, see Bic and Shaw [1988].)

From his enumeration, Lynch [1994] derived some system statistics for a model where resource allocation graphs with the same number of directed edges are equally likely. Our model is quite different. It is more closely related to problems such as linear probe hashing (see Knuth [1973] and Vitter and Flajolet [1990]), computation of random mapping statistics (see Flajolet et al. [1988] and Flajolet and Odlyzko [1990a]), analysis of union-find algorithms (see Yao [1976] and Knuth and Schönhage [1978]), and optimum caching (see Knuth [1985]). These problems are all analyzed in terms of *Abel sums* (so called because many of them can be evaluated explicitly by generalizations of Abel's identity).

The outline of the paper is as follows: In Section 2, we describe our model and derive recurrence relations for various system statistics. In Section 3, we note that all of these recurrences have a common form and give solutions in terms of Abel sums. We prove a result we call the Half and Half Theorem. It says that beginning from any state with j active processes, the expected number of these processes blocked by deadlock time is $(j+1)/2$, one-half more than half the number of initially active processes. In Section 4, we develop a general theory for the evaluation of Abel sums. This is applied to the recurrence solutions of Section 3 to give expressions and inequalities for system statistics. In Section 5, we give asymptotic expressions for these expressions in the case where systems have many processes. In Section 6, we do the same for the case where systems have a fixed number of processes. Here, functions generalizing the Q-function arise. This function was studied by Cauchy [1826], and later by Ramanujan [1912]. Ramanujan actually denoted it θ (as he did several other functions). Knuth was the one to name it Q in his first expected time analysis of an algorithm [Knuth 1968, pages 113–118] (see also Flajolet, et al. [1992]). We use singular expansions to obtain the asymptotics of coefficients of functions of this type.

We observe several notational conventions. The expression $[z^n]s(z)$ denotes the nth coefficient of the generating function $s(z)$. The expression $n^{\underline{m}}$ denotes the falling factorial $n(n-1)\cdots(n-m+1)$. Stacked numbers in braces $\left\{ n \atop m \right\}$ denote Stirling numbers of the second kind (or in the terminology of Graham et al. [1989] the *subset Stirling numbers*). The expression $j!!$ denotes the double factorial function $1 \cdot 3 \cdot 5 \cdots (2j-1)$. $E(\mathbf{X})$ is the expectation of the random variable **X**. We write $f \ll g$ if $f(n)/g(n)$ approaches 0 as $n \to \infty$; this is just another way of writing $f(n) = o(g(n))$.

2. *Recurrence Equations*

In this section, we give a more detailed description of our model and derive recurrence relations for system statistics.

Let m be the number of resources and n be the number of processes in a multiprocessing system. Suppose that the system has just entered a state where there are i free resources and j active processes. Let $\boldsymbol{T}_{i,j}$ be a random variable representing the time to deadlock. $\boldsymbol{T}_{i,j}$ depends on m as well as i and j, but for notational convenience we suppress m. We shall see that $\boldsymbol{T}_{i,j}$ does not depend at all on n.

Consider the arrival time for the next request. If there were just a single active process, this value would be an exponentially distributed random variable with mean 1. For j active processes, it is the minimum of j independent, exponentially distributed random variables with mean 1. A straightforward calculation shows that this is an exponentially distributed random variable with mean j^{-1}. That is, we may express the arrival time for the next request as $j^{-1}\boldsymbol{X}_{i,j}$, where $\boldsymbol{X}_{i,j}$ is an exponentially distributed random variable with mean 1.

Thus, after a time interval of length $j^{-1}\boldsymbol{X}_{i,j}$, a process p chooses a resource r. Now r may be one of the i free resources or one of the $m - i$ allocated resources. The event that r is free is represented by a random variable $\boldsymbol{A}_{i,j}$ where $P[\boldsymbol{A}_{i,j} = 1] = 1 - P[\boldsymbol{A}_{i,j} = 0] = i/m$. When $\boldsymbol{A}_{i,j}$ is 1, an edge is inserted from r to p and the system enters a new state in which there are $i - 1$ free resources and j active processes. When $\boldsymbol{A}_{i,j}$ is 0, we check to see if an edge inserted from p to r would result in deadlock. We do this by constructing a directed path from r as follows: Follow the unique out-edge from r to a process. If this process has out-degree one, again follow the unique out-edge. Continue in this fashion. The resource allocation graph contains no directed cycle yet, so this path terminates. In fact, it terminates at a process, since resources with non-zero in-degree have out-degree one. If the process in which this path terminates is p, an edge from p to r would complete a directed cycle and deadlock results. If the path terminates at any of the other $j - 1$ active processes, we do not have deadlock. Thus, the probability that deadlock does not occur is $(j - 1)/j$. The event that deadlock does not occur may be represented by a random variable $\boldsymbol{B}_{i,j}$ where $P[\boldsymbol{B}_{i,j} = 1] = 1 - P[\boldsymbol{B}_{i,j} = 0] = (j - 1)/j$. Putting this all together, we have

$$\boldsymbol{T}_{i,j} = j^{-1}\boldsymbol{X}_{i,j} + \boldsymbol{A}_{i,j}\boldsymbol{T}_{i-1,j} + (1 - \boldsymbol{A}_{i,j})\boldsymbol{B}_{i,j}\boldsymbol{T}_{i,j-1} \tag{1}$$

where the random variables $\boldsymbol{A}_{i,j}$ and $\boldsymbol{B}_{i,j}$ are independent of each other and of the random variables $\boldsymbol{X}_{i,j}$, $\boldsymbol{T}_{i-1,j}$, and $\boldsymbol{T}_{i,j-1}$.

This model is counterintuitive in one respect. Suppose that process p requests a resource r and at some later time p requests r again. This is not precluded in our description of the model. The system enters deadlock at this point because a 2-cycle appears in the resource allocation graph. If we do not allow the possibility of a process requesting a resource already allocated to it, then the system may never reach deadlock. It may instead reach a state in which all resources have been allocated to a particular process and all other processes are blocked. If we wanted to forbid the possibility of a process requesting a resource already allocated to it, we would have to be more careful about what we mean by expected deadlock time. Of course, the main reason for considering the former model is that it is analytically more tractable.

In the case in which there is a single process, our model is an instance of the famous birthday problem (see Knuth [1968]). This problem asks how many people one must choose at random to find a pair with the same birthday. This is equivalent to labeling 365 resources with the days of the year (we ignore leap days) and computing the expected deadlock time, which, with just one process, occurs only when a resource is repeated. In general, $\boldsymbol{T}_{m,1}$ is the expected number of people one must choose for a year with m days.

Closely related to the problem of finding the expected deadlock time is the problem of finding the total processing time $\boldsymbol{P}_{i,j}$ before deadlock. This requires only a minor modification of the argument used to derive eq. (1). Observe that, if j processes are active for a time interval given by $j^{-1}\boldsymbol{X}_{i,j}$, then the total processing time over that interval is $\boldsymbol{X}_{i,j}$. Hence,

$$\boldsymbol{P}_{i,j} = \boldsymbol{X}_{i,j} + \boldsymbol{A}_{i,j}\boldsymbol{P}_{i-1,j} + (1 - \boldsymbol{A}_{i,j})\boldsymbol{B}_{i,j}\boldsymbol{P}_{i,j-1}. \tag{2}$$

Now if we let $T_{i,j} = E(\boldsymbol{T}_{i,j})$ and $P_{i,j} = E(\boldsymbol{P}_{i,j})$, from (1) and (2) we have, using linearity of expectation and independence, then multiplying by j,

$$jT_{i,j} = 1 + \frac{i}{m}jT_{i-1,j} + \left(1 - \frac{i}{m}\right)(j-1)T_{i,j-1}, \tag{3}$$

$$jP_{i,j} = j + \frac{i}{m}jP_{i-1,j} + \left(1 - \frac{i}{m}\right)(j-1)P_{i,j-1}, \tag{4}$$

for $j > 0$.

A system statistic that will be especially important is $F_{i,j} = P_{i,j}/jT_{i,j}$ which measures system efficiency. Note that $jT_{i,j}$ is total processing time in a system in which j processes are active for a time interval of $T_{i,j}$. The ratio of actual total processing time to this quantity is the expected fraction of time that processes are active.

We will also be interested in the variances of $\boldsymbol{T}_{i,j}$ and $\boldsymbol{P}_{i,j}$. Therefore, we wish to compute the expectations $U_{i,j} = E(\boldsymbol{T}_{i,j}^2)$ and $Q_{i,j} = E(\boldsymbol{P}_{i,j}^2)$. Squaring both sides of (1) and using the identities $\boldsymbol{A}_{i,j}(1 - \boldsymbol{A}_{i,j}) = 0$, $\boldsymbol{A}_{i,j}^2 = \boldsymbol{A}_{i,j}$, $(1 - \boldsymbol{A}_{i,j})^2 = 1 - \boldsymbol{A}_{i,j}$, $\boldsymbol{B}_{i,j}^2 = \boldsymbol{B}_{i,j}$, we have

$$\begin{aligned}\boldsymbol{T}_{i,j}^2 &= j^{-2}\boldsymbol{X}_{i,j}^2 + \boldsymbol{A}_{i,j}\boldsymbol{T}_{i-1,j}^2 + (1 - \boldsymbol{A}_{i,j})\boldsymbol{B}_{i,j}\boldsymbol{T}_{i,j-1}^2 \\ &\quad + 2j^{-1}\boldsymbol{X}_{i,j}\boldsymbol{A}_{i,j}\boldsymbol{T}_{i-1,j} + 2j^{-1}\boldsymbol{X}_{i,j}(1 - \boldsymbol{A}_{i,j})\boldsymbol{B}_{i,j}\boldsymbol{T}_{i,j-1}.\end{aligned}$$

Now take expectations and use eq. (3). Note that $E(\boldsymbol{X}_{i,j}^2) = 2$. Then multiply by j to obtain

$$jU_{i,j} = 2T_{i,j} + \frac{i}{m}jU_{i-1,j} + \left(1 - \frac{i}{m}\right)(j-1)U_{i,j-1}. \tag{5}$$

Performing the same series of computations beginning with (2), we have

$$jQ_{i,j} = 2jP_{i,j} + \frac{i}{m}jQ_{i-1,j} + \left(1 - \frac{i}{m}\right)(j-1)Q_{i,j-1}. \tag{6}$$

Symbol	Description	$S_{i,j}$	$X_{i,j}$	Equation
$T_{i,j}$	Expected Deadlock Time	$jT_{i,j}$	1	(3)
$P_{i,j}$	Expected Total Processing Time	$jP_{i,j}$	j	(4)
$U_{i,j}$	Expected Square of Deadlock Time	$jU_{i,j}$	$2T_{i,j}$	(5)
$Q_{i,j}$	Expected Square of Total Processing Time	$jQ_{i,j}$	$2jP_{i,j}$	(6)
$R_{i,j}$	Expected Number of Resources Allocated	$jR_{i,j}$	ij/m	(7)

FIG. 2. Summary of recurrences.

The last statistic we consider measures how well the system allocates resources. This is also important for system design. Even if the expected deadlock time is fairly long, system performance still might be poor if processes requesting resources are usually blocked. We are interested, therefore, in the expected number of resources $R_{i,j}$ that will be allocated by the time deadlock occurs. This is easy to determine from our model. Notice that whenever a process requests a resource, the probability that the request will be granted is i/m. Thus,

$$jR_{i,j} = \frac{ij}{m} + \frac{i}{m} jR_{i-1,j} + \left(1 - \frac{i}{m}\right)(j-1)R_{i,j-1}. \tag{7}$$

3. *Solving the General Recurrence*

The eqs. (3)–(7) have the form

$$S_{i,j} = X_{i,j} + \frac{i}{m} S_{i-1,j} + \left(1 - \frac{i}{m}\right) S_{i,j-1}, \tag{8}$$

where $S_{i,0} = X_{i,0} = 0$ for all i. (Note that there is no need to specify the values of $S_{0,j}$ since $S_{0,j} = X_{0,j} + S_{0,j-1}$.) We summarize the results of the previous section, with respect to this general form, in Figure 2.

Recurrence (8) is linear in $X_{i,j}$; that is, if $X_{i,j} = aX'_{i,j} + bX''_{i,j}$ for all i and j, then $S_{i,j} = aS'_{i,j} + bS''_{i,j}$, where $S'_{i,j}$ and $S''_{i,j}$ are solutions of the equations obtained from (8) by replacing $X_{i,j}$ with $X'_{i,j}$ and $X''_{i,j}$, respectively.

To solve (8), we might try to form a bivariate generating function $\sum_{i,j} S_{i,j} y^i z^j$ and solve the resulting partial differential equation. Unfortunately, the solution to this equation does not lead easily to a closed-form solution for coefficients of the generating function. Instead, we form sequences of generating functions

$$s_i(z) = \sum_{j=1}^{\infty} S_{i,j} z^j \qquad \text{and} \qquad x_i(z) = \sum_{j=1}^{\infty} X_{i,j} z^j.$$

Multiply eq. (8) by z^j and sum over the range $j \geq 1$ to obtain

$$s_i(z) = x_i(z) + \frac{i}{m} s_{i-1}(z) + \left(1 - \frac{i}{m}\right) z s_i(z).$$

Solving for $s_i(z)$ gives an expression in terms of $s_{i-1}(z)$ and $x_i(z)$:

$$s_i(z) = \frac{m}{m-(m-i)z} x_i(z) + \frac{i}{m-(m-i)z} s_{i-1}(z).$$

Substitute for $s_{i-1}(z)$ on the right side of this expression, then substitute for $s_{i-2}(z)$, and so on. We have

$$s_i(z) = \sum_{k=0}^{i} a_{i,k}(z) x_k(z). \tag{9}$$

where

$$a_{i,k}(z) = \frac{mi(i-1)\cdots(k+1)}{(m-(m-i)z)(m-(m-i+1)z)\cdots(m-(m-k)z)}. \tag{10}$$

It is instructive to use (9) and (10) to compute $T_{m,1}$, which, recall, is the expectation in the birthday problem for a year with m days. Take $X_{i,j} = 1$ for $i \geq 0$ and $j > 0$. Then, $T_{m,1}$ is the linear coefficient of $s_m(z)$ and hence is the sum of the constant coefficients of $a_{m,k}$. Thus,

$$T_{m,1} = 1 + 1 + \frac{m-1}{m} + \frac{(m-1)(m-2)}{m^2} + \cdots + \frac{(m-1)(m-2)\cdots 1}{m^{m-1}}.$$

We may regard this as a function of m. The function $T_{m,1} - 1$ is the Ramanujan Q-function. $T_{m,1}$ is the linear coefficient of $s_m(z)$ and thus is just the first of a family of functions given by the successive coefficients of $s_m(z)$. The techniques used by Knuth and others to obtain the asymptotics of coefficients of the Q-function do not seem to extend easily to other functions in the family. In a later section, we will develop other asymptotic methods to deal with these functions.

Let us return to our analysis of (9) and (10). By partial fraction decomposition,

$$a_{i,k}(z) = \sum_{l=k}^{i} (-1)^{l-k} \binom{i}{i-l,\, l-k,\, k} \left(1 - \frac{l}{m}\right)^{i-k} \left(1 - \left(1 - \frac{l}{m}\right) z\right)^{-1}.$$

Thus,

$$[z^r] a_{i,k}(z) = \sum_{l=k}^{i} (-1)^{l-k} \binom{i}{i-l,\, l-k,\, k} \left(1 - \frac{l}{m}\right)^{i-k+r}$$

and consequently

$$\begin{aligned} S_{i,j} &= \sum_{k=0}^{i} \sum_{r=0}^{j-1} ([z^r] a_{i,k}(z))([z^{j-r}] x_k(z)) \\ &= \sum_{k=0}^{i} \sum_{r=0}^{j-1} \sum_{l=k}^{i} (-1)^{l-k} \binom{i}{i-1,\, l-k,\, k} \left(1 - \frac{l}{m}\right)^{i-k+r} X_{k,j-r} \\ &= \sum_{l=0}^{i} \binom{i}{l} \sum_{k=0}^{l} \binom{l}{k} \left(\frac{l}{m} - 1\right)^{l-k} \sum_{r=0}^{j-1} \left(1 - \frac{l}{m}\right)^{i-l+r} X_{k,j-r} \\ &= \sum_{r=1}^{j} X_{0,r} + \sum_{l=1}^{i} \binom{i}{l} \sum_{k=0}^{l} \binom{l}{k} \left(\frac{l}{m} - 1\right)^{l-k} \sum_{r=0}^{j-1} \left(1 - \frac{l}{m}\right)^{i-l+r} X_{k,j-r}. \end{aligned} \tag{11}$$

This appears a little cumbersome, but for many choices of $X_{i,j}$ there is a considerable simplification. In particular, if $X_{i,j}$ can be written as a product $X_{i,j} = Y_i Z_j$ then we can write

$$S_{i,j} = Y_0 \sum_{r=1}^{j} Z_r + \sum_{l=1}^{i} \binom{i}{l} I_l^{i-1} \sum_{k=0}^{l} \binom{l}{k} Y_k (-I_l)^{l-k} \sum_{r=0}^{j-1} Z_{j-r} I_l^r, \tag{12}$$

where $I_l = 1 - l/m$. The sums $\sum_{k=0}^{l} \binom{l}{k} Y_k (-I_l)^{l-k}$ and $\sum_{r=0}^{j-1} Z_{j-r} I_l^r$ can then be evaluated separately. Results below follow in this way. It is convenient to introduce the notation

$$c_i(p,q) = \sum_{l=1}^{i} \binom{i}{l} \left(\frac{l}{m}\right)^{l-p} \left(1 - \frac{l}{m}\right)^{i-l+q}. \tag{13}$$

Note that $c_i(p+1, q+1) = c_i(p+1, q) - c_i(p,q)$. In the next section, we explain how $c_i(p,q)$ may be computed in closed form for some values of p and q.

THEOREM 3.1. *We have the following expressions for expected deadlock time, expected total processing time, and expected number of resources allocated at deadlock time.*

(*i*) $T_{i,j} = 1 + (c_i(1,0) - c_i(1,j))/j$.
(*ii*) $P_{i,j} = (j+1)/2 + c_i(1,0) - (c_i(2,1) - c_i(2,j+1))/j$.
(*iii*) $R_{i,j} = c_i(1,0) - (c_i(2,1) - c_i(2,j+1))/j$.

PROOF. For (i), take $X_{i,j} = 1$ in eq. (8) so that $Y_i = Z_j = 1$. Use the Binomial Theorem and summation of geometric series to obtain

$$T_{i,j} = 1 + \frac{1}{j} \sum_{l=1}^{i} \binom{i}{l} \left(\frac{l}{m}\right)^{l-1} \left(1 - \frac{l}{m}\right)^{i-1} \left(1 - \left(1 - \frac{l}{m}\right)^{j}\right).$$

Similarly, (ii) and (iii) follow from the identities

$$\sum_{k=0}^{l} \binom{l}{k} k(-I_l)^{l-k} = l(1 - I_l)^{l-1}, \tag{14}$$

$$\sum_{r=0}^{j-1} (j-r) I_l^r = \frac{j}{1 - I_l} - \frac{I_l - I_l^{j+1}}{(1 - I_l)^2}, \tag{15}$$

which are derived by standard techniques. □

Parts (ii) and (iii) of the theorem already provide useful information about system behavior. The expected total processing time $P_{i,j}$ is also the expected number of requests since the request arrival time for each process is 1. The difference between this value and $R_{i,j}$, the expected number of resources allocated, is the expected number of processes blocked by deadlock time. We see that this quantity is always $(j+1)/2$, one half more than half the number of initially active processes. Thus, we have the following surprising result about system performance.

THEOREM 3.2 (THE HALF AND HALF THEOREM). *Beginning from any state with j active processes, the expected number of these processes blocked by deadlock time is* $(j+1)/2$.

There is an easier way to prove the Half and Half Theorem. The theorem is clearly true when $j = 1$. Now suppose the system is in a state where the number j of active processes is more than one and assume the Half and Half Theorem holds for states with $j - 1$ processes. The system may change states a number of times before it changes from a state with j active processes to a state with $j - 1$ active processes. Let us call state changes where the number of active processes decreases *critical*. A system can deadlock only at a critical state change. The probability of deadlock at the first critical state change is $1/j$; in this event, just one of the original processes is blocked. The probability, then, that deadlock does not occur at the first critical state change is $(j - 1)/j$; in this event, the system enters a state with $j - 1$ active processes and the expected number of remaining processes that will be blocked by deadlock times is $j/2$, by the induction hypothesis. Thus, the expected number of processes blocked by deadlock time is $(1/j) + (1 + j/2)(j - 1)/j = (j + 1)/2$.

Using these ideas, we can determine how many processes must be blocked for the probability of deadlock to exceed $\delta > 0$. The probability of reaching deadlock after k critical state changes is

$$\begin{aligned}\frac{1}{j} &+ \frac{j-1}{j}\frac{1}{j-1} + \frac{j-1}{j}\frac{j-2}{j-1}\frac{1}{j-2} + \cdots \\ &+ \frac{j-1}{j}\frac{j-2}{j-1}\cdots\frac{j-k+1}{j-k+2}\frac{1}{j-k+1} = \frac{k}{j}\end{aligned}$$

so we set $k = \delta j$. We have the following result

THEOREM 3.3. *If we allow a system with j initially active processes to operate until δj of those processes are blocked, the probability that deadlock has occurred is δ.*

Before concluding this section, let us explore another consequence of eq. (11). Consider the case where $i = m$, that is, where all resources are initially free. Changing the order of summation we have

$$\begin{aligned}S_{m,j} &= \sum_{r=0}^{j-1}\sum_{k=0}^{m}\binom{m}{k}X_{m-k,j-r}\sum_{l=0}^{k}(-1)^{k-l}\binom{k}{l}\left(\frac{l}{m}\right)^{k+r} \\ &= \sum_{r=0}^{j-1}\sum_{k=0}^{m}\frac{m^{\underline{m-k}}}{m^{m-k}}\left\{{k+r \atop k}\right\}X_{m-k,j-r}.\end{aligned}$$

The last equation follows from formula (6.19) of Graham et al. [1989]. Although this is a more pleasing form than eq. (12) (and shows, in particular, that $S_{m,j}$ is a positive linear combination of the quantities $X_{k,r}$), it is not as useful.

4. *Evaluation of Abel Sums*

Abel's identity is expressed in several different forms in the literature. Here is one of them (see Riordan [1968, page 18]):

$$x\sum_{l=0}^{i}\binom{i}{l}(x+l)^{l-1}(y+i-l)^{i-l} = (x+y+i)^{i}.$$

It follows immediately that

$$\sum_{l=1}^{i}\binom{i}{l}(x+l)^{l-1}(y+i-l)^{i-1}=\frac{(x+y+i)^{i}-(y+i)^{i}}{x}.$$

Taking the limit as x approaches 0 we have

$$\sum_{l=0}^{i-1}\binom{i}{l}l^{l-1}(y+i-l)^{i-l}=i(y+i)^{i-1}.$$

If we put $y=m-i$ and divide by m^{i-1}, we have one of the cases of (13) needed to evaluate the expression in Theorem 3.1(i), viz., $c_i(1,0)=i$. Riordan [1968, pages 18–23] considers several generalizations of Abel's identity. These give some of the other cases $c_i(p,q)$ by the same argument. (N.B. There are some typographical errors in Riordan's formulas.) However, in order to get good asymptotic results, we must develop a more general theory for expressions of the form

$$C_i(p,q,y)=\sum_{l=1}^{i}\binom{i}{l}l^{l-p}(y+i-l)^{i-l+q}.$$

and then evaluate $c_i(p,q)=C_i(p,q,m-i)/m^{i-p+q}$. First note that $C_i(p,q,y)$ is a convolution. Let

$$f_p(z)=\sum_{k=1}^{\infty}\frac{k^{k-p}}{k!}z^k \qquad \text{and} \qquad g_q(z)=\sum_{k=0}^{\infty}\frac{(y+k)^{k+q}}{k!}z^k.$$

(If $p=0$, then it will be convenient to begin the summation for $f_p(z)$ at $k=0$ rather than $k=1$ so that the constant coefficient is 1.) Then $C_i(p,q,y)=[z^i/i!]f_p(z)g_q(z)$. The function $f_1(z)=f(z)$ is well known in combinatorial enumeration. A standard textbook application of the Lagrange Inversion Formula is to show that $f(z)$ is the solution of the functional equation $f(z)=ze^{f(z)}$ (see Wilf [1990]). If we differentiate both sides of this equation, substitute $e^{f(z)}=f(z)/z$, and solve for $f'(z)$ we have $zf'(z)=f(z)/(1-f(z))$. Thus,

$$f_2(z)=\int_0^z\frac{f(x)}{x}\,dx=\int_0^z(1-f(x))f'(x)\,dx=f(z)-\frac{1}{2}f(z)^2,$$

from which it follows that

$$\begin{aligned}f_3(z)&=\int_0^z\frac{f_2(x)}{x}\,dx\\&=\int_0^z\left(1-\frac{1}{2}f(x)\right)(1-f(x))f'(x)\,dx\\&=f(z)-\frac{3}{4}f(z)^2+\frac{1}{6}f(z)^3.\end{aligned}$$

If we use the Lagrange Inversion Theorem to compute the coefficients of $e^{yf(z)}$ and then substitute $f(z)/z$ for $e^{f(z)}$, we have

$$f(z)^y=y\sum_{k=1}^{\infty}\frac{(y+k)^{k-1}}{k!}z^{y+k}.$$

Differentiate, multiply by z, and replace $zf'(z)$ with $f(z)/(1-f(z))$. We obtain

$$g_0(z) = \sum_{k=0}^{\infty} \frac{(y+k)^k}{k!} z^k = \left(\frac{f(z)}{z}\right)^y \frac{1}{1-f(z)}.$$

It follows that

$$\begin{aligned} f_1(z) g_0(z) &= z\left(\frac{f(z)}{z}\right)^{y+1} \frac{1}{1-f(z)} \\ &= \sum_{k=1}^{\infty} \frac{k(y+k)^{k-1}}{k!} z^k, \end{aligned}$$

and so $C_i(1,0,y) = i(y+i)^{i-1}$. A similar evaluation of $f_2(z)g_0(z)$ and $f_3(z)g_0(z)$ shows that

$$\begin{aligned} C_i(2,0,y) &= i(y+i)^{i-1} - \frac{1}{2} i^{\underline{2}}(y+i)^{i-2}, \\ C_i(3,0,y) &= i(y+i)^{i-1} - \frac{3}{4} i^{\underline{2}}(y+i)^{i-2} + \frac{1}{6} i^{\underline{3}}(y+i)^{i-3}. \end{aligned}$$

We can use the method above to compute values of $C_i(p,0,y)$ whenever $p > 0$. We can show easily by induction that

$$f_p(z) = \sum_{k=1}^{p} D_{p,k} f(z)^k.$$

where for all positive p, $D_{p,1} = 1$ and $D_{p+1,k+1} = (D_{p,k+1} - D_{p,k})/k$. These coefficients are the "differences of reciprocals of unity" appearing on page 248 of David et al. [1966]. (This source was located with the help of Sloane [1973]; cf. sequences 2049 and 2305.) It follows that

$$C_i(p,0,y) = \sum_{k=1}^{p} D_{p,i} i^{\underline{k}}(y+i)^{i-k}.$$

We can now compute the values of $c_i(p,q)$ when $p > q \geq 0$. The first few values are as follows.

$$\begin{aligned} c_i(1,0) &= i, \\ c_i(2,0) &= im - \frac{i(i-1)}{2}, \\ c_i(3,0) &= im^2 - \frac{3i(i-1)}{4} m + \frac{i(i-1)(i-2)}{6}, \\ c_i(2,1) &= c_i(2,0) - c_i(1,0) = im - \frac{i(i+1)}{2}, \\ c_i(3,1) &= c_i(3,0) - c_i(2,0) = im^2 - \frac{(3i+1)i}{4} m + \frac{(i+1)i(i-1)}{6}, \\ c_i(3,2) &= c_i(3,1) - c_i(2,1) = im^2 - \frac{(3i+5)i}{4} m + \frac{(i+2)(i+1)i}{6}. \end{aligned}$$

In general,

$$c_i(p,0) = \sum_{k=1}^{p} D_{p,i} i^k m^{i-k}$$

and values of $c_i(p,q)$ for $1 \le q < p$ are obtained by differencing.

Substituting these values for $c_i(p,q)$ into Theorem 3.1, we have the following theorem giving expressions for various system statistics in terms of Abel sums.

THEOREM 4.1. *The expected deadlock time, expected total processing time, and expected number of resources allocated at deadlock time are as follows*:

(*i*) $T_{i,j} = 1 + i/j - c_i(1,j)/j$.
(*ii*) $P_{i,j} = (j+1)/2 + i - im/j + i(i+1)/2j + c_i(2,j+1)/j$.
(*iii*) $R_{i,j} = i - im/j + i(i+1)/2j + c_i(2,j+1)/j$.

From eq. (13), we see immediately that $c_i(1,j) \le c_i(1,0)$ and $c_i(2,j+1) \le c_i(2,1)$, so we have the following important corollary:

COROLLARY 4.2. *The expected deadlock time, expected total processing time, and expected number of resources allocated at deadlock time satisfy the following inequalities*:

(*i*) $1 \le T_{i,j} \le 1 + i/j$.
(*ii*) $(j+1)/2 + i - im/j + i(i+1)/2j \le P_{i,j} \le (j+1)/2 + i$.
(*iii*) $i - im/j + i(i+1)/2j \le R_{i,j} \le i$.

The first of these inequalities suggests that we may want to do deadlock detection at regular intervals rather than at each change of the system state. It shows that even in the worst possible circumstance where resources are never released, expected deadlock time is never less than 1. There is an absolute lower bound for expected deadlock time.

5. *Systems with Many Processes*

The case of systems with many processes is important for applications. This occurs, for example, when i, the number of free resources, is much less than j, the number of active processes. A few of our results require slightly stronger assumptions, either that m, the total number of resources, is much less than j or that $m \log\log m$ is much less than j. The following result is an immediate consequence of Corollary 4.2.

THEOREM 5.1. *Suppose that a system that begins from a state with m resources, i free resources, and j active processes.*

(*i*) (*Expected deadlock time.*) *If* $i = o(j)$, *then* $T_{i,j} \sim 1$.
(*ii*) (*Expected total processing time.*) *If* $i = o(j)$ *and* $m = O(j)$, *then* $P_{i,j} \sim j/2$,
(*iii*) (*Expected number of allocated resources, and system efficiency.*) *If* $m = o(j)$, *then* $R_{i,j} \sim i$, *and* $F_{i,j} = P_{i,j}/jT_{i,j} \sim 1/2$.

We see that a system with many processes performs well; it assigns nearly all resources before deadlock and even though most requests result in a blocked process (since the total number of requests is asymptotically $j/2$ which is much larger than i, the number of resources allocated) system efficiency is still reasonably good. Now let us consider the variances of $T_{i,j}$ and $P_{i,j}$. The proofs are more involved so we state the results separately.

THEOREM 5.2. *In a system that begins from a state with m resources, i free resources, and j active processes, if $m \log\log m = o(j)$, then the variance of deadlock time approaches* 1.

PROOF. We must first obtain bounds on $U_{i,j}$, the mean square of $T_{i,j}$. Recall from Figure 2 that we take $X_{i,j} = 2T_{i,j}$ in eq. (8). Direct substitution gives a very complicated expression, so we substitute instead the lower and upper bounds given by Corollary 4.2.

First, let us derive a lower bound. Let $X_{i,j} = 2$ in eq. (8). Then, we have immediately from Corollary 4.2 that $S_{i,j}/j$ is at least 2.

To obtain an upper bound, we need to solve (8) when $X_{i,j} = 2 + 2i/j$. By linearity, we may solve the cases where $X_{i,j} = 2$ and $X_{i,j} = 2i/j$ separately. In the first case, $S_{i,j}/j$ has an upper bound of $2 + 2i/j$ by Corollary 4.2. In the second case, an explicit solution seems difficult. Instead, we obtain estimates on eq. (12) with $Y_i = 2i$ and $Z_j = 1/j$. By eq. (14), we have

$$S_{i,j} = \sum_{l=1}^{i} \binom{i}{l}\left(\frac{l}{m}\right)^{l-1}\left(1 - \frac{l}{m}\right)^{i-l} \sum_{r=0}^{j-1} \frac{2l}{j-r}\left(1 - \frac{l}{m}\right)^{r}.$$

Recall that

$$\sum_{l=1}^{i} \binom{i}{l}\left(\frac{l}{m}\right)^{l-1}\left(1 - \frac{l}{m}\right)^{i-l} = c_i(1,0) = i.$$

Thus, $S_{i,j}/j$ is a weighted average of the values

$$\frac{2il}{j}\sum_{r=0}^{j-1}\frac{1}{j-r}\left(1 - \frac{l}{m}\right)^{r}. \tag{16}$$

where l ranges from 1 to i. These values can be expressed as $[z^j]g_l(z)$ where

$$g_l(z) = -\frac{2il}{j}\frac{\log(1-z)}{1-(1-l/m)z}.$$

We will show that (16) is $o(1)$ uniformly in l as $j \to \infty$. We do this using ideas related to singularity analysis originally formulated by Darboux (see Henrici [1977] and Flajolet and Odlyzko [1990b]). The dominant singularity of $g_l(z)$ is a logarithmic singularity at $z = 1$ and the only other singularity is at $z = 1 + l/(m - l)$. Apply Cauchy's formula [Markushevich 1977] to obtain

$$[z^j]g_l(z) = -\frac{il}{\pi\sqrt{-1}\,j}\int_\Gamma \frac{\log(1-z)}{(1-(1-l/m)z)z^{j+1}}\,dz. \tag{17}$$

We choose a contour Γ consisting of several parts and bound the integral on each part (see Figure 3). Let δ vary with j so that $0 < -\delta \log \delta \ll 1/j$. Take $\epsilon = 1/(2m)$ and let $\eta > \epsilon$ be a fixed constant less than 1. Γ consists of six pieces $\Gamma_1, \dots, \Gamma_6$. Γ_1 is a circle of radius δ centered at 1 taken clockwise beginning at $1 + \delta$. Γ_2 is a line segment from $1 + \delta$ to $1 + \epsilon$. Γ_3 is a line segment at an angle of $\pi/3$ beginning at $1 + \epsilon$ ending at a point on the circle of radius $1 + \eta$ centered at 0. We obtain Γ_4 by following this circle counterclockwise from this point to its conjugate. Γ_5 is a reflection of Γ_3 through the real axis, but with reversed orientation. Γ_6 is Γ_2 with reversed orientation. In

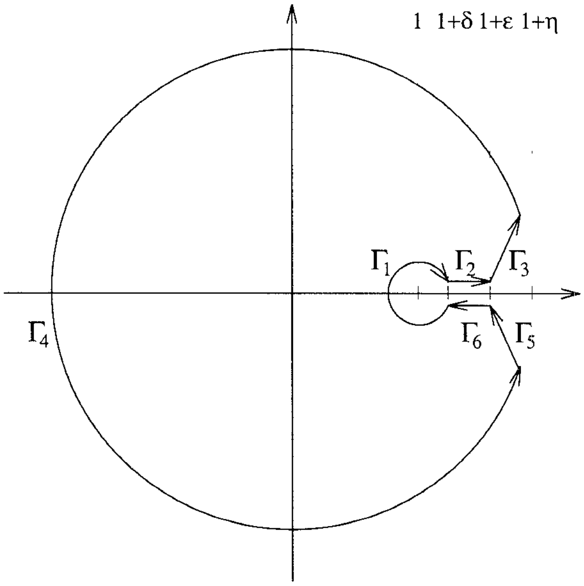

FIG. 3. Contour for evaluation of $U_{l,j}$.

Figure 3, Γ_2 and Γ_6 are shown a small distance from the x-axis for illustrative purposes; they really lie on the x-axis. Note, however that we use the principal branch of $\log z$. That is, the complex plane is slit from 1 to ∞ with Γ_2 and Γ_6 on opposite sides of the slit.

Let us derive bounds on factors of the integrand in (17). First note that Γ was chosen so that it does not come close to the point $1 + l/(m - l)$; the distance is always at least $\sqrt{3}\,l/(4(m - l)) > l/(3(m - l))$ so for all z on Γ

$$\left|\frac{1}{1 - (1 - l/m)z}\right| \le \frac{3m}{l}.$$

For Γ_1, we take $z = 1 - \delta e^{\sqrt{-1}\,\theta}$ for $-\pi \le \theta \le \pi$. Thus, $\log(1 - z) = \log \delta + \sqrt{-1}\,\theta$ and so $|\log(1 - z)| \le -\log \delta + \pi$. Also, $1 - \delta \le |z|$ so $1/|z^{j+1}| \le 1/(1 - \delta)^{j+1}$. Thus, (17) is bounded in modulus by

$$\frac{6im\delta(-\log \delta + \pi)}{j(1 - \delta)^{j+1}},$$

which approaches 0 as j increases since $-\delta \log \delta \ll 1/j$.

Next combine the contributions to (17) from Γ_2 and Γ_6. These two contours are the same, except for orientation, but the branches of $\log(1 - z)$ used differ by $-2\pi\sqrt{-1}$. Therefore, the total contribution is

$$\frac{2il}{j}\int_{1+\delta}^{1+\epsilon} \frac{dx}{(1 - (1 - l/m)x)x^{j+1}} \le \frac{6im}{j}\int_{1+\delta}^{1+\epsilon} \frac{dx}{x^{j+1}} \sim \frac{6im}{j^2},$$

which approaches 0 as j increases.

Next, for Γ_3 we take $z = 1 + \epsilon + we^{\sqrt{-1}\,\pi/3}$ where w ranges from 0 to a value slightly less than $2(\eta - \epsilon)$—the exact value is not crucial.

We claim that $-\log(1-z) = -\log(-\epsilon - w\exp(\sqrt{-1}\,\pi/3))$ attains its maximum modulus on Γ_3 when $w = 0$. Consider the function $F(x, y) = |-\log(x + \sqrt{-1}\,y)|^2$. The directional derivative of F in the direction given by the unit vector $(-1/2, -\sqrt{3}/2)$ (which is in the direction of the contour) is

$$\frac{1}{r^2}\left(\frac{x}{2} + \frac{\sqrt{3}\,y}{2}\right)\log\left(\frac{1}{r^2}\right) - (\sqrt{3}\,x - y)\theta,$$

where $x + \sqrt{-1}\,y = r\exp(\sqrt{-1}\,\theta)$. If $\theta < 0$, $r < 1$, and $y \le \sqrt{3}x \le 0$, this quantity will be negative. Notice that these conditions are satisfied by $x + \sqrt{-1}\,y = -\epsilon - we^{\sqrt{-1}\,\pi/3}$ when $0 \le w \le 2(\eta - \epsilon)$. We see that the modulus of $-\log(1-z)$ decreases along the contour and, thus, attains a maximum value of $\sqrt{(\log\epsilon)^2 + \pi^2} \le \log(2m) + \pi$. Notice also that $|z| \ge 1 + \epsilon + x/2$ on Γ_3 so the contribution is bounded by

$$\frac{6im(\log(2m) + \pi)}{j}\int_0^\infty \frac{dx}{(1 + \epsilon + x/2)^{j+1}} = \frac{12im\log m}{j^2}\left(1 - \frac{1}{2m-1}\right)^j$$
$$\le \frac{12im\log m}{j^2}e^{-j/(2m-1)},$$

which approaches 0 since $m\log\log m \ll j$. We obtain a bound on Γ_5 in exactly the same way.

Finally, on Γ_4 we see that the modulus of the denominator of the integrand grows exponentially in j and the numerator is bounded, so the contribution to (17) approaches 0.

We conclude that $U_{i,j} \sim 2$ and hence that the variance of $\boldsymbol{T}_{i,j}$ approaches 1. □

The $\log\log m$ factor in the preceding theorem is somewhat annoying. It seems likely that it can be eliminated. Note that the only place in the proof where we used this condition, rather than the weaker assumption that $m = o(j)$, was in bounding the integral on Γ_3 and Γ_5. We could instead have used the condition that $i < m/\log m$.

THEOREM 5.3. *In a system that begins from a state with m resources, i free resources, and j active processes, if $i = o(j)$ and $m = O(j)$, then the variance of total processing times is asymptotic to $j^2/12$.*

PROOF. We begin similarly to the proof of the previous theorem. We obtain bounds on $Q_{i,j}$, the mean square of $\boldsymbol{P}_{i,j}$ by substituting the lower and upper bounds given by Corollary 4.2. For the lower bound we take $X_{i,j} = j(j+1) + 2ij - 2im + i(i+1)$ in eq. (8) and for the upper bound we take $X_{i,j} = j(j+1) + 2ij$. Thus, by linearity, we have four cases to consider.

Case 1. $X_{i,j} = j(j+1)$. In (12), we let $Y_i = 1$ and $Z_j = j(j+1)$. Use the identities

$$\sum_{r=1}^{j} r(r+1) = \frac{j(j+1)(j+2)}{3},$$

and

$$\sum_{r=0}^{j-1}(j-r)(j-r+1)I_l^r = \frac{j(j+1)}{1-I_l} - 2\frac{jI_l}{(1-I_l)^2} + 2\frac{I_l^2 - I_l^{j+2}}{(1-I_l)^3}$$

to obtain

$$\frac{S_{i,j}}{j} = \frac{(j+1)(j+2)}{3} + (j+1)c_i(1,0) - 2c_i(2,1) + 2(c_i(3,2) - \frac{c_i(3,j+2))}{j}.$$

Case 2. $X_{i,j} = 2ij$. In (12), we let $Y_i = 2i$ and $Z_j = j$. We have from (14) and (15) that

$$\frac{S_{i,j}}{j} = \frac{2mc_i(1,0) - 2m(c_i(2,1) - c_i(2,j+1))}{j}.$$

Case 3. $X_{i,j} = -2im$. In (12), we let $Y_i = -2im$ and $Z_j = 1$. We have

$$\frac{S_{i,j}}{j} = -\frac{2m^2(c_i(1,0) - c_i(1,j))}{j}.$$

Case 4. $X_{i,j} = i(i+1)$. In (12), we let $Y_i = i(i+1)$ and $Z_j = 1$. Use the identity

$$\sum_{k=0}^{l} \binom{l}{k} k(k+1)(-I_l)^{l-k} = l(l-1)(1-I_l)^{l-2} + 2l(1-I_l)^{l-1},$$

to obtain

$$\frac{S_{i,j}}{j} = \frac{(m^2+m)(c_i(1,0) - c_i(1,j))}{j} - \frac{m(c_i(2,0) - c_i(2,j))}{j}.$$

A lower bound for $S_{i,j}/j$ in case 1 is

$$\frac{(j+1)(j+2)}{3} + i(j+1) - 2im + i(i+1).$$

Add the values of $S_{i,j}/j$ for cases 2, 3, and 4. Using the identities for $c_i(p,q)$ developed in Section 3 we have

$$\frac{2im - 3m(c_i(2,1) - c_i(2,j+1))}{j} - \frac{(m^2-m)(c_i(1,0) - c_i(1,j))}{j}.$$

A lower bound for this is

$$\frac{2im - 4im^2}{j} + \frac{i(3i+5)m}{2j}.$$

Thus, taking all the cases together we have a lower bound for $Q_{i,j}$ of

$$\frac{(j+1)(j+2)}{3} + i(j+1) + i(i+1) - \frac{4im^2}{j} + \frac{i(3i+5)m}{2j}.$$

To obtain an upper bound for $Q_{i,j}$, add the values of $S_{i,j}/j$ in cases 1 and 2. This is less than

$$\frac{(j+1)(j+2)}{3} + i(j+1) + i(i+1) + \frac{2im^2}{j} - \frac{i(3i+5)m}{2j} + \frac{(i+2)(i+1)i}{3j}.$$

We see that under the hypotheses of the theorem that $Q_{i,j} \sim j^2/3$. The square of the mean of $\boldsymbol{P}_{i,j}$ is asymptotic to $j^2/4$ so the variance is asymptotic to $j^2/12$. $\square$

These two results on the variance of $\boldsymbol{T}_{i,j}$ and $\boldsymbol{P}_{i,j}$ show that we must be cautious if we consider doing deadlock detection only at regular intervals in systems with many processes. Even though deadlock time and total processing time have reasonable means, their standard deviations are constant multiples of their means. For design purposes, we would like to have more information about the distributions of these random variables. We have not yet succeeded in determining the distributions explicitly. Perhaps simulations would give some insight. We will make some brief comments about the distribution in the last section.

6. *Systems with a Fixed Number of Processes*

Systems with a fixed number of processes are also important for applications. We will need more sophisticated asymptotic techniques than in the last section to carry out the analysis of these systems. We will also need to restrict to the case $i = m$, that is, the case where all resources are initially free. By Theorem 4.1, we see that we need to determine asymptotic expansions of $c_m(1, j) = [z^m/m!](f(z)f_{-j}(z)/m^{m+j-1})$ and $c_m(2, j+1) = [z^m/m!](f_2(z)f_{-j-1}(z)/m^{m+j-1})$.

We first need to find the singular expansions of the functions f_p. The singular expansion of $f_1(z) = f(z)$ can be derived without too much difficulty from the Implicit Function Theorem (see Markushevich [1977]) and the defining equation $f(z) = ze^f(z)$. Set $g(z,w) = w - ze^w$ and note that $w = f(z)$ is the solution of $g(z,w) = 0$. Now $g(1/e, 1) = (\partial g/\partial w)(1/e, 1) = 0$ and $(\partial^2 g/\partial w^2)(1/e, 1) \neq 0$, so f has an algebraic singularity of order $1/2$ at $z = 1/e$. Let $\delta = (2 - 2ez)^{1/2}$ (the $2^{1/2}$ factor is to simplify succeeding expressions) so $z = (2 - \delta^2)/2e$. Substituting into the equation $g(z,w) = 0$ we have $\delta = (2 - 2we^{1-w})^{1/2}$. That is, f, considered as a function of δ, satisfies $\delta = (2 - 2f(\delta)e^{1-f(\delta)})^{1/2}$. In this context, f is the functional inverse of $h(w) = (2 - 2we^{1-w})^{1/2}$. It is a simple matter to compute the Taylor expansion of h and then of its inverse (in Maple, the *reversion* operation accomplishes this). We obtain

$$f(z) = 1 - \delta + \frac{1}{3}\delta^2 - \frac{11}{72}\delta^3 + \frac{43}{540}\delta^4 - \frac{769}{17280}\delta^5 + \cdots.$$

A similar derivation of this formula can be found in Flajolet and Odlyzko [1990a]. The singular expansion of $f_2(z) = f(z) - f(z)^2/2$ can now be seen to be

$$f_2(z) = \frac{1}{2} - \frac{1}{2}\delta^2 + \frac{1}{3}\delta^3 - \frac{5}{24}\delta^4 + \frac{47}{360}\delta^5 - \cdots.$$

Note that application of the operator $z(d/dz)$ to $f_j(z)$ gives $f_{j-1}(z)$. Now $z(d/dz)\delta^m = -m\delta^{m-2} + m\delta^m/2$ so

$$f_0(z) = \delta^{-1} + -\frac{2}{3} - \frac{1}{24}\delta + \frac{2}{135}\delta^2 - \frac{23}{3456}\delta^3 + \cdots,$$

$$f_{-1}(z) = \delta^{-3} - \frac{11}{24}\delta^{-1} - \frac{4}{135} - \frac{1}{1152}\delta - \cdots.$$

In general,

$$f_{-j}(z) = \sum_{k=-2j-1}^{\infty} b_{j,k}\delta^k$$

where $b_{j,k} = kb_{j-1,k}/2 - (k+2)b_{j-1,k+2}$. We can solve this recurrence first for the coefficients $b_{j,-2j-1}$, then for the coefficients $b_{j,-2j}$ and so on. For $j \geq 2$, we have

$$f_{-j}(z) = j!!\Bigg(\delta^{-2j-1} - \frac{12j^2-1}{24(2j-1)}\delta^{-2j+1} + \frac{144j^4 - 384j^3 + 264j^2 - 23}{1152(2j-1)(2j-3)}\delta^{-2j+3} + \cdots\Bigg).$$

Thus,

$$f(z)f_{-1}(z) = \delta^{-3} - \delta^{-2} - \frac{1}{8}\delta^{-1} + \frac{149}{540} - \frac{767}{17280}\delta + \cdots,$$

$$f(z)f_{-j}(z) = j!!\Bigg(\delta^{-2j-1} - \delta^{-2j} - \frac{12j^2-16j+7}{24(2j-1)}\delta^{-2j+1} + \frac{18j^2-11j+4}{36(2j-1)}\delta^{-2j+2} + \frac{2160j^4 - 11520j^3 + 18104j^2 - 10528j + 3063}{17280(2j-1)(2j-3)}\delta^{-2j+3} + \cdots\Bigg)$$

for $j \geq 2$, and

$$f_2(z)f_{-j-1}(z) = (j+1)!!\Bigg(\frac{1}{2}\delta^{-2j-3} - \frac{12j^2+72j+35}{48(2j+1)}\delta^{-2j-1} + \frac{1}{3}\delta^{-2j} + \frac{144j^4 + 1344j^3 - 216j^2 - 144j - 47}{2304(2j+1)(2j-1)}\delta^{-2j+1} + \cdots\Bigg).$$

for $j \geq 1$. By the Binomial Theorem and Stirling's formula $[z^m/m!]\delta^{-\alpha}$ is

$$\frac{\pi^{1/2}}{\Gamma(\alpha/2)}\frac{m^{m+(\alpha-1)/2}}{2^{(\alpha-1)/2}}\Bigg(1 + \frac{3\alpha^2 - 6\alpha + 2}{24m} + \frac{9\alpha^4 - 60\alpha^3 + 120\alpha^2 - 72\alpha + 4}{1152m^2} + O\left(\frac{1}{m^3}\right)\Bigg).$$

Substituting this into the expressions above and simplifying, we have

$$c_m(1,1) = m - \frac{(2\pi)^{1/2}}{2} m^{1/2} + \frac{1}{3} - \frac{(2\pi)^{1/2}}{24} m^{-1/2} + \frac{4}{135} m^{-1} + O(m^{-3/2})$$

$$c_m(1,j) = m - \frac{2^{1/2}\Gamma(j+1/2)}{\Gamma(j)} m^{1/2} + \frac{2j-1}{3} - \frac{2^{1/2}\Gamma(j+1/2)}{36\Gamma(j)} \frac{4j^2-6j+5}{2j-1} m^{-1/2} - \frac{2j^2-4j-6}{135} m^{-1} + O(m^{-3/2})$$

$$c_m(2, j+1) = \frac{1}{2} m^2 - \frac{2j+1}{2} m + \frac{2^{1/2}\Gamma(j+1/2)}{\Gamma(j)} \frac{2j+1}{3} m^{1/2} - \frac{j^2}{3} + O(m^{-1/2})$$

The asymptotic formula for $c_m(1,1)$ agrees with eq. (25) in 1.2.11.3 of Knuth [1968]. Our approach seems somewhat simpler since it does not require analysis of the incomplete gamma function. From Theorem 4.1, we now have the main result of this section.

THEOREM 6.1. *Fix j. For systems beginning from a state with j processes and all resources free, we have the following asymptotic formulas for expected deadlock time, expected total processing time, expected number of resources allocated at deadlock time, and system efficiency.*

$$T_{m,1} = \frac{(2\pi)^{1/2}}{2} m^{1/2} + \frac{2}{3} + \frac{(2\pi)^{1/2}}{24} m^{-1/2} - \frac{4}{135} m^{-1} + O(m^{-3/2}),$$

$$T_{m,j} = \frac{2^{1/2}\Gamma(j+1/2)}{\Gamma(j+1)} m^{1/2} + \frac{j+1}{3j} + \frac{2^{1/2}(4j^2-6j+5)\Gamma(j-1/2)}{72\Gamma(j+1)} m^{-1/2} + \frac{2j^2-4j-6}{135j} m^{-1} + O(m^{-3/2}), \qquad \text{when } j \geq 2$$

$$P_{m,j} = \frac{2^{3/2}\Gamma(j+3/2)}{3\Gamma(j+1)} m^{1/2} + \frac{j+3}{6} + O(m^{-1/2}),$$

$$R_{m,j} = \frac{2^{3/2}\Gamma(j+3/2)}{3\Gamma(j+1)} m^{1/2} - \frac{j}{3} + O(m^{-1/2}),$$

$$F_{m,j} = \frac{2}{3} + \frac{1}{3j} + O(m^{-1/2}).$$

Thus, we find that for a fixed number of processes, deadlock time, total processing time, and number of resources allocated are asymptotic to a constant multiple of $m^{1/2}$. System efficiency always approaches a value greater than $2/3$. This seems very promising. Unfortunately, we have not yet determined the order of the variances for deadlock time and total processing time and we have no idea of the distribution.

7. *Conclusions*

Our results suggest that it may be worthwhile to investigate doing deadlock detection at regular intervals in some systems. The most compelling evidence for this is the absolute lower bound for expected deadlock time (Corollary 4.2(i)). Also, the high values for system efficiency for systems with many processes (Theorem 5.1(iii)) and systems with a fixed number of processes (Theorem 6.1) are encouraging.

If we knew the distribution of deadlock time $\boldsymbol{T}_{i,j}$ we could determine, for a given $\delta > 0$, the length of time the system must operate for the probability of deadlock to be δ. We could then compare the cost of continuously updating to do deadlock detection versus the expected cost of resolution and rollback when doing deadlock detection at regular intervals. Even though we do not know the distribution, Theorem 3.3 provides a similar, but less satisfactory, approach. We can count blocked processes. This requires some communication between processes and may therefore cost more than waiting for a fixed time interval, but it is still likely to cost less than deadlock prevention.

What do we have to do to determine the distribution of deadlock time $\boldsymbol{T}_{i,j}$ as i or j increase? A standard approach for determining the distribution of a limit random variable is to find its characteristic function (see Feller [1971]). From eq. (1), we have the following recurrence for the characteristic function $\hat{\boldsymbol{T}}_{i,j}$ of $\boldsymbol{T}_{i,j}$:

$$\hat{\boldsymbol{T}}_{i,j} = \frac{1}{1 - \sqrt{-1}\,z/j}\left(\frac{i}{m}\hat{\boldsymbol{T}}_{i-1,j} + \left(1 - \frac{i}{m}\right)\left(1 - \frac{1}{j}\right)\hat{\boldsymbol{T}}_{i,j-1} + \left(1 - \frac{i}{m}\right)\frac{1}{j}\right).$$

This may lead to a solution, but it does not seem to work well with the generating function methods used here.

There are other important questions still to be answered.

First, the absolute lower bound we derived for expected deadlock time, although theoretically optimal, certainly does not come near what occurs in practice. The most obvious reason for this is that processes do release resources in multiprocessing systems so deadlock time is usually much greater than in the model here. We need to develop a more realistic model, but this will not be easy. It will require adding a queuing-theoretic dimension to the combinatorial problem of determining when cycles emerge in certain kinds of random bipartite graphs.

Second, we need to deal with the problem of 2-cycles mentioned in Section 2. We would expect processes to remember the resources allocated to them and to avoid making duplicate requests. As we noted, if we prohibit 2-cycles, then it is no longer clear what is meant by expected deadlock time. It is also not clear how to modify the model. Should we proceed as before with a process randomly choosing a resource but have the system to continue in case a 2-cycle arises? Or should we require a process to choose randomly among resources not allocated to it? In either case, the model becomes much more complicated. We must keep track not only of the number of active processes, but also of the number of resources allocated to each active process.

In addition to the cases in Sections 4 and 5, we would like to have asymptotic expressions for system statistics when $i = Kj$. We would like to know in Section 4 if conditions like $m = o(j)$ and $m \log\log m = o(j)$ can be replaced with $i = o(j)$. We would like to solve the general case in Section 5, not just the case $i = m$.

Finally, we would like to know the variances and distributions of $\boldsymbol{T}_{i,j}$ and $\boldsymbol{P}_{i,j}$ for fixed j. There are some unpublished results on certain cases of this problem. In the case $j = 1$, variance and distribution can be determined. This problem should not be confused with one related to the problem of determining the variance and distribution for the traditional birthday problem. Even though the mean of $\boldsymbol{T}_{m,1}$ is the expected number of people needed to find a pair with the same birthday for m day years, variance behaves differently. $\boldsymbol{T}_{m,1}$ models the "street corner birthday problem": A person stands on a busy street corner and asks passers-by their birthdays until a repetition is found. Arrival times of passers-by are exponentially distributed random variables (with normalized mean 1). In the traditional birthday problem (the "classroom birthday problem"), a teacher in a classroom asks students their birthdays until a repetition is found. Arrival times are precisely 1. Clearly, the variance for the street corner birthday problem is greater than for the classroom birthday problem. Donald Knuth (personal communication) has informed us that using techniques in Knuth [1985] one may show that variance in this case is asymptotic to $(4 - \pi)m/2$. Guy Louchard of the Universite Libre de Bruxelles (personal communication) has informed us that, for both the classroom and street corner birthday problems, the random variable for a repeated birthday, suitably normalized, converges in distribution to a Rayleigh distributed random variable. Alfredo Viola of the University of Waterloo (personal communication) has some preliminary results on the variance of $T_{m,j}$ for fixed j.

ACKNOWLEDGMENTS. We would like to thank Philippe Flajolet for many useful suggestions and discussions. We would also like to thank Yuri Gurevich and Andreas Blass for making a fundamental observation that led to the development of the model in eq. (1). Both authors would like to thank the creators of Maple system. Most of the computations in this paper were carried out or checked in Maple.

REFERENCES

BIC, L., AND SHAW, A. C. 1988. *The Logical Design of Operating Systems*. Prentice-Hall, Englewood Cliffs, N.J.

CAUCHY, A. L. 1826. *Exercises de Mathematique*. Chez de Bure Freres, Paris, France.

DAVID, F. N., KENDALL, M. G., AND BARTON, D. E. 1966. *Symmetric Function and Allied Tables*. Cambridge University Press, Cambridge, England.

FELLER, W. 1971. *Introduction to Probability Theory and Its Applications*, vol. 2. Wiley, New York.

FLAJOLET, P., GARDY, D., AND THIMONIER, L. 1988. Probabilistic languages and random allocations. In *Automata, Languages and Programming*, T. Lepistö and A. Salomaa, eds. Lecture Notes in Computer Science, vol. 317. Springer-Verlag, pp. 239–253.

FLAJOLET, P., GRABNER, P. J., KIRSCHENHOFER, P., AND PRODINGER, H. 1992. On Ramanujan's Q-function. Rapport de Recherche 1760. Institut National de Recherche en Informatique et en Automatique, Rocquencourt, France.

FLAJOLET, P., AND ODLYZKO, A. M. 1990a. Random mapping statistics. In *Advances in Cryptology*, J. J. Quisquater and J. Vandewalle, eds. Lecture Notes in Computer Science, vol. 434. Springer-Verlag, New York, pp. 329–354.

FLAJOLET, P., AND ODLYZKO, A. M. 1990b. Singularity analysis of generating functions. *SIAM J. Disc. Math. 3*, 216–240.

GRAHAM, R. L., KNUTH, D. E., AND PATASHNIK, O. 1989. *Concrete Mathematics*. Addison-Wesley, Reading, Mass.

HENRICI, P. 1977. *Applied and Computational Complex Analysis*, vol. 2. Wiley, New York.

KNUTH, D. E. 1968. *The Art of Computer Programming*, vol. 1: *Fundamental Algorithms*, 2nd ed. Addison-Wesley, Reading, Mass.

KNUTH, D. E. 1973. *The Art of Computer Programming*, vol. 3: *Sorting and Searching*. Addison-Wesley, Reading, Mass.

KNUTH, D. E. 1985. An analysis of optimum caching. *J. Algorithms 6*, 181–199.

KNUTH, D. E., AND SCHÖNHAGE, A. 1978. The expected linearity of a simple equivalence algorithm. *Theoret. Comput. Sci. 6*, 281–315.

LYNCH, J. 1994. Random resource allocation graphs and the probability of deadlock. *SIAM J. Disc. Math. 7*, 458–473.

MARKUSHEVICH, A. I. 1977. *Theory of Functions of a Complex Variable*. Chelsea, New York. Revised. (English edition translated and edited by Richard A. Silverman.)

RAMANUJAN, S. 1912. Question 294. *J. Indian Math. Soc. 4*, 151–152.

RIORDAN, J., 1968. *Combinatorial Identities*. Wiley, New York.

SLOANE, N. J. A. 1973. *Handbook of Integer Sequences*. Academic Press, San Diego, Calif.

VITTER, J. S., AND FLAJOLET, P. 1990. Analysis of algorithms and data structures. In *Handbook of Theoretical Computer Science*, vol. A: *Algorithms and Complexity*, J. van Leeuwen, ed. North Holland, Amsterdam, The Netherlands, pp. 431–524.

WILF, H. S. 1990. *generatingfunctionology*. Academic Press, San Diego, Calif.

YAO, A. C.-C. 1976. On the average behavior of set merging algorithms. In *Proceedings of the 8th ACM Symposium on Theory of Computing* (Hershey, Pa., May 3–5). ACM, New York, pp. 192–195.

RECEIVED JANUARY 1993; REVISED SEPTEMBER 1994; ACCEPTED NOVEMBER 1994