Data structures maxima

Guy Louchard, Claire Kenyon, René Schott

UNITÉ DE RECHERCHE
INRIA-LORRAINE

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P. 105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

N° 1394

Programme 2
Calcul symbolique, Programmation et Génie logiciel

DATA STRUCTURES MAXIMA

Guy LOUCHARD
Claire KENYON
René SCHOTT

Février 1991

Data Structures Maxima
Maxima des structures de données

Guy Louchard*, Claire Kenyon† and René Schott‡

Abstract

The purpose of this paper is to analyze the maxima properties (value and position) of some data structures. Our theorems concern the *distribution* of the random variables. Previously known results usually dealt with the *mean* and sometimes the variance of these random variables. Many of our results rely on diffusion techniques. That is a very powerful tool, which has already been used with some success in the analysis of algorithms.

Résumé

Dans cet article nous étudions la taille maximale atteinte par certaines structures de données dynamiques au bout d'un temps assez long. Nous explicitons complètement les distributions limites des variables aléatoires et améliorons de ce fait notablement les résultats connus jusqu'à présent qui concernaient la moyenne et parfois la variance de ces variables aléatoires. Les techniques de base sont celles de la théorie des diffusions.

*Laboratoire d'Informatique Théorique, Université Libre de Bruxelles, CP 212, Boulevard du Triomphe, 1050 Bruxelles, Belgium

†LIENS, URA CNRS 1327, ENS, 45 rue d'Ulm, 75230 Paris Cedex 05, France

‡CRIN, INRIA-Lorraine, Université de Nancy 1, 54506- Vandoeuvre-lès-Nancy, France

Data Structures Maxima

G. Louchard[*], Claire Kenyon[†] and R. Schott[‡]

Abstract

The purpose of this paper is to analyze the maxima properties (value and position) of some data structures. Our theorems concern the *distribution* of the random variables. Previously known results usually dealt with the *mean* and sometimes the variance of these random variables. Many of our results rely on diffusion techniques. That is a very powerful tool, which has already been used with some success in the analysis of algorithms.

1 Introduction

This paper concerns the maximum size reached by a dynamic data structure over a long period of time. The notion of "maximum" is basic to resource preallocation. The expected value of the maximum size had already been studied in Kenyon-Mathieu and Vitter [28] with elementary methods. Our goal here is two-fold:

1. Derive more precise asymptotic expressions for the expected maximum (with lower order terms).

2. When possible, find the whole distribution of the maximum size.

Our proofs involve advanced analytic and probabilistic techniques; in particular, they use Laplace transforms, complex analysis around singularities, diffusion processes and Brownian motions. They rely on some results which are well-known in the world of probabilists. Our hope is to show that diffusion techniques are a powerful tool for the average-case analysis of algorithms.

Given a dynamic data structure, whose size evolves through time with each new insertion or deletion, the main requirement for formalizing the maximum-size problem in a mathematical framework

[*]Laboratoire d'Informatique Théorique, Université Libre de Bruxelles, CP 212, Boulevard du Triomphe, 1050 Bruxelles, Belgium

[†]LIENS, URA CNRS 1327, ENS, 45 rue d'Ulm, 75230 Paris Cedex 05, France

[‡]CRIN, INRIA-Lorraine, Université de Nancy 1, 54506- Vandoeuvre-lès-Nancy, France

is to define the distribution of the sequence of arrivals (insertions) and departures (deletions). The models we consider come from three main sources.

1. Probabilistic model: that is the world of queueing theory, whose assumptions have been well tested by time. We study the $M/M/1$, $M/G/1$, $G/M/1$, $M/M/\infty$ and $G/G/\infty$ queueing systems. For example, in the $M/M/\infty$ model, the number of servers is infinite, so that a client newly arrived is served right away: there is no waiting time. The interarrivals and the service times are all independent and follow an exponential distribution.

2. Combinatorial models: Françon introduced in the 70's the concept of file histories, which are beautiful combinatorial objects [14]. A file history is a labelled path, where each elementary step of the path is of the type $(x, y) \rightarrow (x+1, y+/-1)$ or $(x+1, y)$. The x-axis represents time, and the y-axis represents the size of the file. The relative weights of the paths depend on the type of file history considered (stack, linear list, dictionary, symbol table, or priority file). However, we must say that not many real-life situations correspond to the distribution of file histories.

3. Hashing with lazy deletion. In that non-Markovian data structure, introduced by Van Wyk and Vitter [44], the data structure is a separate-chaining hash table, and the arrivals and lifetimes follow a process of type 1 or 2 above, but the items are not removed from the table as soon as they "die": instead, we wait until there is a new insertion in a chain of the table before removing the "dead" items from that chain. This enables to save on the access cost to the table.

The results are as follows.

In the first part, we study the maximum size reached by a queue over $[0 \ldots t]$ when t is goes to infinity, for the most classical types of queues.

When there is only one server processing the requests of the arriving customers, we find expressions describing the distribution for an $M/M/1$, $M/G/1$ or $G/M/1$ process. The technique uses Laplace transforms and developments in the neighborhood of a singularity.

When there is an infinite number of servers, there is no waiting time, and a client only stays in the system as long as it takes for him to be served. This has been studied in [38], and, in [28], there is an equivalent to the expected value of the maximum size in the $M/M/\infty$ case. In this paper, through a different approach we get more precise expressions on the expectation. We also solve the $G/G/\infty$ case. Our proof involves reducing the problem to an Ornstein-Uhlenbeck diffusion process, whose properties are well-known (by probabilists).

In the second part, we study some specific kinds of dynamic data structures.

First, we look at combinatorial objects defined by [14] and by Knuth for modelling the evolution of dictionaries, stacks, linear lists, priority files and symbol tables: file histories. The average value of the maximum was evaluated in [28], who obtained an equivalent. Using sophisticated results on Brownian Bridges [11], we show how to get a much more precise estimate of the maximum, with several lower-order terms. As an example, we analyze priority queues in the standard model and dictionaries in Knuth's model.

Second, we study the maximum size of hashing with lazy deletion (a time-saving non-Markovian variant of hash tables, introduced in [44]). In that paper, various models of distribution were suggested. By repeated use of Daniels' theorem [11], we get precise estimates of the average maximum in all the models, with lower order terms (which are not obtainable by the methods of [28]).

Finally, we look at the limiting profiles of a file history, i.e. what fraction of time is spent by the data structure at a given level l. Again we show that we can obtain the asymptotic distributions.

2 The maximum size of a single server queue

In this section, we study several classical queueing theory models, $M/M/1$, $M/G/1$ and $G/M/1$, to find the evolution of the maximum size reached over a (long) period of time. The average value of the maximum is already known, but for the purpose of resource preallocation, it is much more useful to know at what level l the probability that the maximum size reaches l becomes small: then it is sufficient to reserve space l to the data structure. In this section, we find the *entire distribution* of the maximum size as $t \to \infty$. The proofs of this section are based on a few ideas:

1. Instead of the maximum size, study the hitting time, i.e. the length of time before the data structure first reaches size l. (It has the advantage of having a continuous distribution).

2. Take Laplace transforms to obtain closed (complicated) formulas for the transform of the hitting time (with the help of Cohen [7]).

3. Approximate the Laplace transform for levels $l \to \infty$. In all cases, we find that the Laplace transform, as a function of s, is equivalent to $1/(1 - s/s^*)$, where s^* is the critical singularity of the function. Thus the hitting time is asymptotically distributed as exponential random variable.

4. When (as in the $M/G/1$ and $G/M/1$ cases) s^* does not appear in closed form but as the l^{th} coefficient of an analytic function: approximate that function around its smallest singularity to get an equivalent for s^*.

$Max(t)$ will denote the maximum size reached over time interval $[0 \dots t]$: thus $Max(t) = \max_{t' \in [0,t]} Q(t')$, where $Q(t')$ is the queue length. Our problem is finding the distribution of $Max(t)$. To solve this kind of problems, two approaches are usually available: either deal directly with the hitting time distribution, or compute the extreme value distribution of the maximum on a busy period: then we use renewal theory (with the busy cycle mean). However, as the queue length is a *discrete* random variable, the latter approach cannot be used; we will see that actually, we do not obtain a *unique* limit distribution for $M(t)$ but rather a family of distributions depending on the fractional part of $C_1 \log(t) + C_2$, where C_1, C_2 are some constants.

This situation is rather similar to other data structures used in computer science : see for instance the number of registers in arithmetic expression evaluation (Flajolet and Prodinger [15], Louchard [30]), the binary trie (Louchard [30], Flajolet and Steyaert [16]), the digital search tree (Louchard

[31]), approximate counting (Flajolet [17]). The extreme-value distribution, $e^{-e^{-x}}$ frequently appears in this context. We will first analyze the $M/M/1$ queue, then the $M/G/1$ and $G/M/1$ queues. We could not obtain the $G/G/1$ limit theorem.

2.1 The M/M/1 case

In this section, we study the simplest of all single-server queues: the $M/M/1$ queue. In this model, the customers arrive in the system in such a way that the interarrival times are independent and exponentially distributed, with mean $1/\lambda$. When a customer arrives, its request starts being processed by the server right away if there are no other customers present. Otherwise the new customer must wait in line. The service times (for successive requests) are assumed to be independent and exponentially distributed with mean $1/\mu$. Thus the average number of customers present in the stationary case is $\rho = \lambda/\mu$, which we will always assume is smaller than 1.

We want to know the evolution of the maximum size reached over a (long) period of time. The average value of that maximum is already known, but for the purpose of resource preallocation, it is much more useful to know at what level j the probability that the maximum size reaches j becomes small: then it is sufficient to reserve space j for the structure. In the following theorem, we find *the entire distribution* of the maximum size as $t \to \infty$.

Theorem 1 *Let t and j go to infinity in such a way that $t = \Theta(1/\rho^j)$. Then we have:*

$$\Pr\{Max(t) \le j\} \sim e^{-\rho^j t\lambda(1-\rho)^2}$$

and:

$$\Pr\{Max(t) = j\} \sim e^{-\rho^j t\lambda(1-\rho)^2} - e^{-\rho^j t\lambda(1-\rho)^2/\rho}.$$

Thus we see that the distribution of $Max(t)$ is closely related to $exp(-exp(-j))$ when t is large. We also observe that if $f(u,j) = \Pr\{Max(e^u) \le j\}$, then f is periodic in u, with period $\log(\rho^{-1})$, in the sense that

$$f(u + \log(\rho^{-1}), j) = f(u, j-1).$$

Proof:

From now on, we will assume that $\mu = 1$ (without loss of generality, since the general case can be deduced through a simple change of time scale $t \to \mu t$).

The maximum of the process over $[0 \dots t]$ is $\le j$ if and only if the size of the process first reaches $j+1$ at a time $\ge t$. Thus we will study the distribution of T_l, the hitting time for size l (with $l = j+1$). This random variable has the advantage of having a continuous distribution. Let $\phi_l(s)$ be the Laplace transform of T_l, $\phi_l(s) = E(e^{-sT_l})$ with $s \ge 0$. There is an exact formula for $\phi_l(s)$:

from [7], equation II.2.47, we know that $\phi_l(s)$ is the ratio of two Laplace transforms of transition probabilities, and in [7], I.4.28, we find an explicit form for those Laplace transforms. Thus:

$$\phi_l(s) = \frac{\rho^l x_2^l/(\rho(x_1 - x_2)) + \rho^l x_2^l(1 - \rho x_2)^2}{1/(\rho(x_1 - x_2)) + \rho^l x_2^{2l}(1 - \rho x_2)^2},$$

where x_2 is the smaller root of the equation

$$\rho X^2 - (1 + \rho + s)X + 1 = 0.$$

The asymptotic behaviour of T_l is related to the negative singularity s_1^* of $\phi_l(s)$ with smallest absolute value (in fact, $s_1^* \to 0$ as $l \to \infty$). Developing $\phi_l(s)$ when $l \to \infty$, we obtain:

$$\phi_l(s) \sim \frac{1}{1 - \frac{s}{s_1^*}} \text{ when } s < s_1^* \text{ and } l \to \infty,$$

with $s_1^* := -\rho^l(1 - \rho)^2$. Taking the inverse Laplace transform, we get

$$\Pr\{T_l \geq t\} \sim e^{-\rho^l(1-\rho)^2 t}$$

as $l \to \infty$ and $t = \Theta(1/\rho^l)$. Thus

$$\Pr\{Max(t) \leq j\} = \Pr\{T_{j+1} \geq t\} \sim e^{-\rho^{j+1}(1-\rho)^2 t}.$$

In the general case, where $\mu \neq 1$, the corresponding formula is:

$$\Pr\{Max(t) \leq j\} \sim e^{-\rho^j \rho(1-\rho)^2 \mu t},$$

which proves the first part of the theorem.

The second part is a direct consequence. □

We noted that the distribution of $Max(t)$ is "periodic" in $\log t$. The asymptotic moments of $Max(t)$ are also periodic functions of $\log t$. The asymptotic non-periodic term in the moments are given by the first term in the Fourier expansion.

Theorem 2 *The constant term $\bar{E}$ in the Fourier expansion, in $\psi_1(t) = \log(\lambda(1-\rho)^2 t)/\log(\rho^{-1})$, of the moments of $Max(t)$ is asymptotically given by*

$$\bar{E}\{M(t) - \psi_1(t)\}^i \sim \int_{-\infty}^{+\infty} x^i e^{-\rho^x} dx$$

The extreme-value distribution function $e^{-e^{-x}}$ is well known and has mean γ and variance $\pi^2/6$ (Johnson and Kotz [26] p.272). From this, we can for instance derive

$$\bar{E}(Max(t)) \sim \psi_1(t) + \frac{1}{2} + \frac{\gamma}{\log(\rho^{-1})}$$

2.2 The G/M/1 case

In the $G/M/1$ case, as in the $M/M/1$ case, there is only one server, and the successive service times are independent and have an exponential distribution of mean $1/\mu$; the interarrival times are still independent, of mean $1/\lambda$, but their distribution is now general. Let $A(t)$ be the distribution function: $A(t) = \Pr\{\text{interarrival time} \le t\}$. Let $\alpha(s)$ denote the Laplace transform of $A(t)$:

$$\alpha(s) = \int_0^\infty e^{-st} dA(t).$$

Using an approach similar to the $M/M/1$ case, we can also describe the distribution of $Max(t)$, the maximum queue size over $[0 \ldots t]$.

Theorem 3 *Let $\tilde{z}$ denote the smallest root of the equation $z = \alpha((1-z)\mu)$. Then we have as j and $t \to \infty$ with $t = \Theta(1/\tilde{z}^j)$:*

$$\Pr\{Max(t) \le j\} \sim e^{\tilde{z}^j t\lambda(1-\tilde{z})B_1},$$

with the constant B_1 defined by

$$B_1 = \int_0^\infty \mu s e^{-(1-\tilde{z})\mu s} dA(s) - 1.$$

As in the previous section, there is still a periodicity property and an expression for the first term of the Fourier expansion of $Max(t)$.

We can check that Theorem 1 is just a special case of Theorem 3: if $A(t)$ is an exponential distribution, $A(t) = 1 - e^{-\lambda t}$, then it is easy to see that $\alpha(s) = 1/(1 + s/\lambda)$, hence $\tilde{z} = \lambda/\mu = \rho$ and $B_1 = \rho - 1$.

Proof:

We study the distribution of the hitting time τ_l for size l *before* an arrival. Let $\phi_l(s)$ denote its Laplace transform. From [7], II.3.36-39, we get an exact formula for $\phi_l(s)$:

$$\phi_l(s) = \frac{1}{1 - \left[\frac{1-\alpha(s)}{(1-z)[z-\alpha(s+(1-z)\mu)]}\right]_{l-1}},$$

where $[f(z)]_l$ denotes the coefficient of z^l in the development of $f(z)$. We now find a approximation of $\phi_l(s)$ as $l \to \infty$:

$$\phi_l(s) \sim \frac{1}{1 - \frac{s}{s_2^*}},$$

with

$$s_2^* := \frac{\lambda}{\left[\frac{1}{(1-z)(z-\alpha((1-z)\mu))}\right]_{l-1}}.$$

But $T_l \equiv \tau_{l-1}$. Thus, as $l \to \infty$, we have:

$$\Pr\{T_l \geq t\} \sim e^{\lambda t/[1/(1-z)(z-\alpha((1-z)\mu)]_{l-2}}.$$

So far, our proof follows the lines of the proof of Theorem 2.1. Now we need an equivalent of

$$\left[\frac{1}{(1-z)(z-\alpha((1-z)\mu))}\right]_{l-2}.$$

To this end, we will approximate the bracketed quantity in the neighborhood of its smallest singularity. From [7], II.3.16, we find that $z - \alpha((1-z)\mu)$ has a unique root, say $\tilde{z}$, inside the unit circle. We get

$$\frac{1}{(1-z)(z-\alpha((1-z)\mu)} \sim \frac{1}{(1-\tilde{z})B_1(\tilde{z}-z)}$$

as $z \to \tilde{z}$, with $B_1 = \int_0^\infty exp(-(1-\tilde{z})\mu t)\mu t dA(t) - 1$. Therefore

$$\left[\frac{1}{(1-z)(z-\alpha((1-z)\mu))}\right]_{l-2} \sim \frac{1}{(1-\tilde{z})B_1\tilde{z}^{l-1}},$$

and

$$\Pr\{Max(t) \leq j\} = \Pr\{T_{j+1} \geq t\} \sim e^{\lambda t(1-\tilde{z})B_1\tilde{z}^j},$$

and the theorem is proved.

Remark 1

In [7], III.Th.7.5, is given an upper and lower bound for the distribution of $M(n)$, the maximum number of customers in n busy cycles. First of all, it is easily checked that the quantities ψ and c_1 used in that theorem are given here by $\psi \equiv (1-\tilde{z})\mu$, $c_1 \equiv -B_2$. But, by [7] II.3.49 we know that the busy cycle mean length is given by $\bar{\ell} := 1/[\lambda(1-\tilde{z})]$. From renewal theory, we know that $t/n(t) \to \bar{\ell}, t \to \infty$ ($n(t)$ is the number of busy cycles in $[0,t]$). If we could use the second approach described in the beginning of this section, we could easily derive our theorem from Cohen's approach (with his formula on top of p.625). Our result would actually correspond to the lower hand of [7],Th.7.5. Similarly, in the $M/M/1$ case, our theorem 2.1 corresponds to the lim inf given in Anderson [3] (example on p.112).

Remark 2

In Iglehart [24], for the $G/G/1$ queue, a (positive) quantity γ is defined by $E[e^{\gamma x}] = 1$ where $x := u_a - u_b$ ($u_a :=$ interarrival time, $u_b :=$ service time). It is easy to check that, here, $\alpha(\gamma)/(1-\gamma/\mu) = 1$, that $\gamma \equiv \psi$ (see remark 1) $\equiv (1-\tilde{z})\mu$. Then, the waiting times limit-distributions of Theorems 7.2 and 7.5 of [7] are immediate from Iglehart's Theorem 2.

Remark 3

When $\rho \to 1$ in some particular sense, limit distributions are available : see Serfozo [42] and [43] for such results.

Remark 4

In a recent report, Sadowsky and Szpankowski [40] analyse the $G/G/s$ case. (In our case $s = 1$). They prove, under some regularity conditions, that the maximum number of customers on a busy cycle has a geometric tail, with parameter $\alpha(\gamma)$ (γ as given in remark 3). This off course only leads to an upper and lower bound for the distribution of $Max(t)$.

2.3 The M/G/1 case

In the $M/G/1$ model, the interarrivals are independent and have an exponential distribution of mean $1/\lambda$, the service times are independent but have a general distribution $B(t)$ of mean $1/\mu$.

In that case, the calculations are much more complicated. The proof goes along the same lines as in the previous section, and makes use of Cohen's results [7].

Let $\beta(s) = \int_0^\infty e^{-st} dB(t)$ be the Laplace transform of $B(t)$.

Theorem 4 *If $s_0 < 0$, s_0 being the abscissa of convergence of $\beta(s)$, and if $\beta(s) \underset{s \to s_0}{\to} \infty$, then let $\tilde{y}$ be the root > 1, nearest to 1 of $z = \beta((1-z)\lambda)$. Then, for $j \to \infty$:*

$$\Pr\{Max(t) \leq j\} \sim e^{\tilde{y}^{-j} t \lambda (1-\rho)^2 (\tilde{y}-1)/B_2}$$

if $t = \Theta(\tilde{y}^{-j})$, with

$$B_2 = 1 - \int_0^\infty e^{-(1-\tilde{y})\lambda t} \lambda t dB(t).$$

Theorem 2 is still applicable if we replace $\log \rho$ by $-\log \tilde{y}$.

Proof: The departure times now lead to an imbedded Markov chain, where the states correspond to the number of customers just *after* a departure. Once again, we consider the Laplace transform $\varphi_j(s)$ of the hitting time into state j. From [7] II.4.35, we have:

$$\varphi_j(s) = [y(s)]^{-j} \left(1 - (1 + \frac{s}{\lambda} - y(s)) \frac{\left[\frac{\beta(s+(1-z)\lambda)}{\beta(s+(1-z)\lambda)-z}\right]_j}{\left[(1 + s/\lambda - z) \frac{\beta(s+(1-z)\lambda)}{\beta(s+(1-z)\lambda)-z}\right]_j} \right),$$

where $y(s)$ is the root of $z = \beta(s + (1 - z)\lambda)$ with smallest absolute value. From [7] II.4.40, we find an approximate value of $y(s)$: $y(s) = 1 - \frac{s}{\mu(1-\rho)} + o(s)$. The rest of the proof consists in expanding the formula giving $\phi_j(s)$ in the neighborhood of $s = 0$ and as $j \rightarrow \infty$. After various asymptotic analyses, we get

$$\varphi_j(s) \sim 1/[1 - s/s_3^*], \qquad s \rightarrow s_3^*, j \rightarrow \infty$$

where $s_3^* := \lambda(1 - \rho)^2(\tilde{y} - 1)/[\tilde{y}^j B_2]$. The theorem follows from there.

3 The infinite-server queue maximum problem

In the models where there is an infinite number of servers, a customer arriving always finds a server ready to process its requests. For example, that can model a dynamic data structure, with items being inserted in the structure (i.e. customers arriving), staying in for a certain length of time (i.e. service time) and finally being deleted and removed (i.e. exiting the system after service). We will assume that the interarrivals are independent, as well as the service times, and that their distribution is known.

3.1 The M/M/ ∞ model

In the $M/M/\infty$ model, the arrivals follow a Poisson distribution of parameter λ, and the service times have an exponential distribution of average $1/\mu$. Here again, we are interested in the maximum number of clients in the system during $[0 \ldots T]$, where $T \rightarrow \infty$: that would tell us whether a real system (in which the number of servers is not really infinite) has a high risk of becoming one day overloaded.

There are already known results about the average value of the maximum [28]. Here, we derive much more precise results concerning the exact distribution of the maximum, and lower order terms on the expectation.

We will assume that λ, μ and T all go to infinity in such a way that $\lambda/\mu \rightarrow \infty$ (i.e. the average number of items present in the data structure is large) and $\mu T \rightarrow \infty$ (i.e. we look at the dynamic structure during a time interval long enough so that a lot of changes occur). Our results are valid in the range $\ln(\lambda T) = o(\lambda/\mu)$ only, but that covers most realistic applications; the other cases are presently under study.

Up to doing a time-scale change, we can assume that instead of looking at the maximum of a (λ, μ)-process over $[0 \ldots T]$, we are looking at the maximum of a $(\lambda T, \mu T)$-process over $[0 \ldots 1]$. Thus, if we define $\lambda' = \lambda T$ and $\mu' = \mu T$, and if $Q_{(\lambda',\mu')}(t')$ denotes the number of customers in the system at time t', we want the distribution of

$$\max_{t' \in [0 \ldots 1]} Q_{(\lambda', \mu')}(t').$$

We will do another time-scale change, so as to get within the frame of a useful theorem by Iglehart [21]. Let $t = \mu' t'$. Then $\lambda'' = \lambda'/\mu' = n$, and $\mu'' = 1$. Then we know from Iglehart [21] that

$$\frac{Q_{(n,1)}(t) - n}{\sqrt{n}} \Rightarrow OU(t),$$

as n goes to infinity and t is large[1]. $OU(t)$ denotes the Ornstein-Uhlenbeck process (which describes the Brownian motion of a harmonically bound particle), and has infinitesimal mean -1, infinitesimal variance 2 and covariance $e^{-(t-s)}$ for $t \geq s$. From the error term in Iglehart [23] and Feller [13] (XVI.7) we see that the convergence is correct as far as $[OU(t)]^3 = o(\sqrt{n})$.

Let $\mu' = 1/h(n) \to \infty$. We want to study

$$\max_{t \in [0, 1/h(n)]} Q_{(n,1)}(t).$$

Since

$$Q_{(n,1)}(t) \sim n + \sqrt{n} OU(t),$$

(t large, n going to infinity), we want to investigate

$$\max_{t \in [0, 1/h(n)]} Q_{(n,1)}(t) \sim n + \sqrt{n} \max_{t \in [0, 1/h(n)]} OU(t).$$

We will now use a powerfull theorem by Berman [4]. Let $m_t = Max_{s \in [0 \ldots t]} OU(s)$. Then from [4] we can prove that

$$\Pr\{u(t)(m_t - u(t)) \leq v\} \sim e^{-e^{-v}}$$

as t goes to infinity, where $u(t)$ satisfies

$$u^2 \frac{t e^{-u^2/2}}{\sqrt{2\pi}} \sim 1.$$

After some algebra, we obtain

$$u(t) \sim \sqrt{2 \ln t} + \frac{\ln \ln t - \ln \pi}{2\sqrt{2 \ln t}} + o(\frac{1}{\sqrt{\ln t}}).$$

[1] Here $\Rightarrow$ denotes the weak convergence of random functions in the space of all right-continuous functions having left limits and endowed with the Skorohod metric (see Billingsley [6])

Thus

$$\Pr\{m_t \le x\} \sim e^{-e^{-u(t)(x-u(t))}}.$$

The distribution $exp(-exp(-v))$ has mean γ, the Euler constant. Thus we get the distribution of m_t and hence of our $M/M/\infty$ process.

Theorem 5 *Let $Q_{(\lambda,\mu)}(t)$ denote the size at time t of an $M/M/\infty$ process with parameters λ and μ. If $\ln(\lambda t) = o(\lambda/\mu)$, let $\ln(\lambda t) = \frac{\lambda}{\mu} g(\frac{\lambda}{\mu})$ with $g(n) \underset{n\to\infty}{\to} 0$. If $g(n) = o(n^{-2/3})$, then we have*

$$E(Max_{s\in[0\ldots t]} Q_{(\lambda,\mu)}(s)) \sim \frac{\lambda}{\mu} + \sqrt{\frac{\lambda}{\mu}}\left[\sqrt{2\ln(\mu t)} + \frac{\ln\ln(\mu t)}{2\sqrt{2\ln(\mu t)}} + \frac{\gamma - \ln \pi/2}{\sqrt{2\ln(\mu t)}} + o\left(\frac{1}{\sqrt{\ln(\mu t)}}\right)\right].$$

The theorem can also be derived using Keilson's result [27] or Jaeschke's approach [25], but Berman's approach is the only one which generalizes to the $G/G/\infty$ case. Other cases are the object of work in progress. For instance, if $g(n) = \xi n^{-2/3}$, it appears that we must add $(2\xi)^{3/2}/6$ to γ.

3.2 The G/M/ ∞ model

In this case the interarrivals are independent but have a general distribution of mean $1/\lambda$. Let σ^2 be the variance of the interarrival times. Iglehart's weak convergence theorem generalizes [22], so that

$$\frac{\sqrt{2}(Q_n(t) - n)}{\sqrt{n(1+\sigma^2)}} \Rightarrow OU(t), \ t \text{ large}.$$

As before we obtain:

Theorem 6 *If $\lambda, \mu, t \to \infty$ in such a way that $\mu t \to \infty, \rho \to \infty$, and $\log(\lambda t) = o(\rho)$, then, if $g(n) = o(n^{-2/3})$,*

$$E(\max_{[0t]} Q_{\lambda,\mu}(s)) \sim \frac{\lambda}{\mu} + \sqrt{\frac{1+\sigma^2}{2}}\sqrt{\frac{\lambda}{\mu}}$$

$$\left[\sqrt{2\ln(\mu t)} + \frac{\ln\ln(\mu t)}{2\sqrt{2\ln(\mu t)}} + \frac{\gamma - \ln \pi/2}{\sqrt{2\ln(\mu t)}} + \left(\frac{1}{\sqrt{\ln(\mu t)}}\right)\right].$$

3.3 The G/G/ ∞ case

In this even more general case, the interarrivals follow a general distribution of mean $1/\lambda$ and variance σ^2, and the service times are also independent of general distribution $B(t) = \Pr\{\text{service time} \leq t\}$. We now use another generalization of Iglehart [23] (the limiting process is a non-Markovian stationary Gaussian process) and results of Berman [4] (16.9). After some algebra, we deduce the following result

Theorem 7 *Assume that $B(t)$ has a continuous density, and that λ, μ, t satisfy the same conditions as in the $G/M/\infty$ case above. Let*

$$C_2 = 1 + (\sigma^2 - 1)\int_0^\infty (1 - B(t))^2 dt.$$

Then

$$E(\max_{[0t]} Q_{\lambda,\mu}(s)) \sim \frac{\lambda}{\mu} + \sqrt{C_2\frac{\lambda}{\mu}}$$

$$\left[\sqrt{2\ln(\mu t)} + \frac{\ln\ln(\mu t)}{2\sqrt{2\ln(\mu t)}} + \frac{\gamma - \ln\pi/2}{\sqrt{2\ln(\mu t)}} + o\left(\frac{1}{\sqrt{\ln(\mu t)}}\right)\right].$$

$$+\log(\frac{1+\sigma^2}{2C_2})\frac{\sqrt{C_2\frac{\lambda}{\mu}}}{\sqrt{2\ln(\mu t)}}.$$

4 List structures maxima

In this section, we leave the world of queueing theory models to enter that of combinatorial models for dynamic data structures. More specifically, we are interested in file histories as invented by Françon and developped by Flajolet, Françon and Vuillemin. The size of a dynamic data structure increases by 1 with each insertion and decreases by 1 with each deletion. The successive sizes, as a function of time, form a path on the plane, which starts at level 0 (the data structure is initially empty) and normally returns to 0 after a sequence of n operations (insertions, deletions, or queries). A probability distribution is defined by weighing the paths according to the data structure under study. File histories have been studied by Françon, Flajolet, Puech, Vuillemin and Viennot in particular, and they have discovered beautiful links with orthogonal polynomials, continued fractions and various combinatorial objects. In [32] [33], Louchard, Schott and Randrianarimanana have developed a complete probabilistic analysis of these structures. In [28], the average value of the height of the path (maximum size) was given for some kinds of file histories. In this section, once again we look at the distribution of the maximum size. Assume that the operations happen at times 1,2,...,n. Our technique is based on the observation that the process can be decomposed into two simple components: let $S(t)$ denote the size of the data structure at time t ($1 \leq t \leq n$). Then $S(t) = \tilde{y} + Z(t)$, where $\tilde{y}$, the *average* size of the structure at time t, is a fairly simple curve (for instance a concave parabola in our first subsection), and $Z(t)$ is a (small) Gaussian process.

4.1 Daniels' fundamental result

All the results in the rest of the paper are based on a general theorem by Daniels [11]. We want information about $Max_{t\in[0...1]}Y(t)$, where Y is a certain random process. Assume that $Y(t)$ can be written as

$$Y(t) = \tilde{z}(t) + Z(t),$$

where $\tilde{z}(t)$ is a certain deterministic curve and $Z(t)$ is a random Gaussian process, of covariance $C(s,t)$. Note that $\tilde{z}(t)$ is *not* random. Let M be the maximum of $Y(t)$ for $t \in [0...1]$, and t^* be the first time at which the maximum M is reached. At that time, $Z(t^*) = M - \tilde{z}(t^*)$, and that is the first time that Z reaches that value. Thus we can look for the hitting time of $Z(t)$ to the absorbing boundary $M - \tilde{z}(t)$. Near that crossing point, $Z(t)$ locally behaves like a Brownian motion (or a variant of it, such as a Brownian bridge) Durbin [12]. It is also known that the hitting time and place densities for a B.B. can be deduced from the hitting time density for a B.M. (see for instance Louchard [33] for a constant boundary and Csaki et al [9] for a general proof).

Suppose that we are looking at the maximum size of a data structure over time interval [0...1] when some parameter n goes to infinity (for instance, n might be the number of operations). We will assume that $\tilde{z}(t)$ satisfies

$$\tilde{z}(t) = \sqrt{n} z(t),$$

where $z(t)$ is independent of n (that assumption will be true for all the applications in this paper). Moreover, assume that $z(t)$ has a unique maximum, reached at time $\bar{t}$, with $z(\bar{t}) = 0$ (up to doing a translation we can always assume that).

Daniels has matched the local behaviour of $C(s,t)$ with the Brownian bridge covariance near $\bar{t}$ [11]. From [11] and Daniels and Skyrme [10], we can deduce information about the maximum M and the time t^* when it is reached

Notations

Let A and B be the constants defined by

$$A = [\partial_s C]_{\bar{t}} + |[\partial_t C]_{\bar{t}}| \text{ and } B = (-z''(\bar{t}))^{-1/3}.$$

Let

$$G(x) = \frac{2^{-1/3}}{2\pi i} \int_{-i\infty}^{i\infty} e^{sx} \frac{ds}{A_i(-2^{1/3}s)},$$

where A_i is the classical Airy function, and let λ be the universal constant defined by

$$\lambda = \int_{-\infty}^{\infty} [e^{x^3/6} G(x) - max(x,0)]dx \sim 0.99615.$$

Let

$$u = n^{1/3} A^{-1/3} B^{-2} (t^* - \bar{t}).$$

Theorem 8 *With the above assumptions and notations, if* $[\partial_t C(s,t)]_{\bar{t}} \leq 0$ *and* $[\partial_s C(s,t)]_{\bar{t}} > 0$*, then we have:*

1. *M is asymptotically Gaussian with mean and variance*

$$\begin{cases} E(M) &= \lambda B n^{-1/6} A^{2/3} + O(n^{-1/3}). \\ \sigma^2(M) &= C(\bar{t},\bar{t}) + O(n^{-1/3}). \end{cases}$$

2. *The conditional maximum $M|t^*$ is asymptotically Gaussian with mean and variance*

$$\begin{cases} E(M|t^*) &= n^{-1/6} A^{-1/3} B[[\partial_s C]_{\bar{t}} \frac{G'(-u)}{G(-u)} + |[\partial_t C]_{\bar{t}}| \frac{G'(u)}{G(u)}] + O(n^{-1/3}). \\ \sigma^2(M|t^*) &= C(\bar{t},\bar{t}) + O(n^{-1/3}). \end{cases}$$

3. *The joined density of M and t^* is given by*

$$\phi(M,t) = \frac{2}{\sqrt{2\pi C(\bar{t},\bar{t})}} e^{-M^2/2C(\bar{t},\bar{t})} \left\{ G(t)G(-t) + n^{-1/6} B A^{-1/3} \frac{M}{C(\bar{t},\bar{t})} \right.$$

$$\left. [[\partial_s C]_{\bar{t}} G(t) G'(-t) + |[\partial_t C]|_{\bar{t}} G(-t) G'(t)] + O(n^{-1/3}) \right\}.$$

4. *u has density* $2G(u)G(-u)[1 + O(n^{-1/3})]$.

Proof:
A direct new proof is given in the Appendix.

4.2 File histories

A file history represents the evolution of the size of a dynamic data structure by a path of length n =number of operations performed (insertions, queries or deletions), going from level 0 to level 0. In [28], the average value of the height of the path (=maximum size) was given for some kinds of data structures, but only a rough equivalent was found.

Here we study the example of priority file histories [18]. Other structures such as linear lists or dictionaries could be analyzed with similar techniques.

A priority list is by definition a data structure on which no queries are performed: only insertions and deletions occur, and moreover, deletions happen only for the minimum. Thus, when a new item is inserted in a priority list of size k, there are $(k+1)$ intervals defined by the elements already present, and to which the new item may belong, but when an item is deleted, there is only one possible choice. This is reflected in the weights of the paths, see [18].

Assume that there are $2n$ operations performed during the history, n insertions and n deletions. Let $Y(t)$ be the size of the data structure after the $\lfloor nt \rfloor^{th}$ operations $(0 \leq t \leq 2)$. ¿From Louchard [32], we know that

$$\forall t \quad \frac{Y(t) - \frac{1}{2}nt(2-t)}{\sqrt{n}} \Rightarrow X(t) \text{ when } n \to \infty,$$

where $X(t)$ is a Markovian Gaussian process with mean 0 and covariance

$$C(s,t) = \frac{s^2(2-t)^2}{4} \text{ when } s \leq t.$$

The error term can be deduced for the various expansion in [32] Sec.4. It appears that the relative error in the density is $O(\frac{1}{\sqrt{n}})$ (non uniform in X). Thus, in this case, using the notations of Daniels' theorem, we have $z(t) = t(2-t)/2 - 1/2$, with maximum at $\bar{t} = 1$ and $z''(\bar{t}) = -1$ Since the covariance here is such a simple function, we can easily apply the theorem, and we find:

Theorem 9 *If $Y_n(t)$ denotes the size of a random priority file history of length $2n$ at time $\lfloor nt \rfloor$, then we have:*

$$E(Max_t Y_n(t)) = \frac{n}{2} + \lambda n^{1/3} + O(n^{1/6}),$$

and more generally

$$Max_t Y_n(t) = \frac{n}{2} + \sqrt{n}\, M + O(n^{1/6})$$

is reached at time t^*, with M and t^* given by Daniels' theorem. We can prove that the error therm in the weak convergence to $X(t)$ is negligible *with the error in $MaxY$.*

Actually, this is not the first example of applications of diffusion processes to computer science. Other structures have been analyzed using brownian excursions or brownian meandering; see [30] for height in planar trees and stack structures maxima.

4.3 List structures in Knuth's model of file histories

This statistical model has been introduced by Knuth [29] and Françon [19]. We refer to [20], [36] for a detailed presentation and the analysis of dynamic algorithms in this model. As an example, we will analyse the dictionary. Let us just remember that if the size of this data structure is k, then the number of possibilities is equal to k for a deletion or a successful query, but that the number of possibilities for the i-th insertion or unsuccessful query is equal to i. The application of Daniels' theorem is much more technical. From Louchard, Schott and Randrianarimanana [33] Theorem 9, it appears that the size Y of the dictionary satisfies

$$\frac{Y([nt]) - nz(t)}{\sqrt{n}} \Rightarrow X(t) + Z(t), \qquad 0 \le t \le 2$$

where $z(t) := [\sqrt{\frac{3}{2}}\,(2\cos(\theta/3) - 1)/2]^{1/2}$, with $\theta := \arccos[2t(2-t) - 1]$.
$X(\cdot)$ is a Gaussian Markovian process with mean 0 and covariance

$$CX(s,t) = \gamma(s)\,\gamma(2-t), \qquad s \le t$$

with $\gamma(s) := [3 + z'(s)][z'(s) - 1]^3/8$. We have $z'(\bar{t}) = 0$ and $\bar{t} = 1$.
$Z(\cdot)$ is a Gaussian non-Markovian process, given by

$$Z(t) = \int_o^2 \Psi(t,u)\sqrt{s(z'(u))}\,dBM(u)$$

where BM$(\cdot)$ is a classical Brownian Motion and $s(z') := (1 - z'^2)/8$.

$$\psi(t,u) = \begin{cases} \gamma(2-t)[\frac{\gamma(u)}{z(u)} + \gamma'(u)\frac{\log(z(u))}{2s(z')}\frac{\partial s}{\partial z'}] & \text{if } t \ge u \\ \gamma(t)[\frac{\gamma(2-u)}{z(u)} - \gamma'(2-u)\frac{\log(z(u))}{2s(z')}\frac{\partial s}{\partial z'}] & \text{otherwise} \end{cases}$$

The covariance of Z can be written as

$$CZ(s,t) = \int_o^2 \psi(s,u)\psi(t,u)s[z'(u)]du$$

At $\bar{t} = 1$, we derive

$$CZ(1,1) = 2\gamma^2(1)\int_o^1 \left[\frac{\gamma(u)}{z(u)} - \gamma'(u)\frac{\log(z(u))}{2s(z')}\frac{z'(u)}{4}\right]^2 s(z')du$$

We have been unable to obtain an explicit expression for this quantity. A numerical evaluation gives $CZ(1,1) = .013992....$ After a detailed analysis, it appears that $cZ_1 = [\partial s CZ(s,t)]_{1,1} = cZ_2 = 0$. Let us now turn to the properties of $X(\cdot)$. We check that $\gamma(1) = -3/8$, $CX(1,1) = 9/64$. From [33], Theorem 7, we have $z(1) = 3/8, z''(1) = -2/3$. It is now easy to deduce

$$[\partial_s C]_{\bar{t}} = 1/4,\ [\partial_t C]_{\bar{t}} = -1/4,\ A = 1/2,\ B = (2/3)^{-1/3},\ C(\bar{t},\bar{t}) = 0.154617...$$

This finally leads to

Theorem 10

$$Max := \max Y([nt]) \sim \frac{3}{8}n + \sqrt{n}\,M + 0(n^{1/6})$$

and the maximum is reached at time t^*, where M and t^* are given by Daniels' theorem.

5 Limiting profiles of list structures

In the last section, we analyzed the distribution of the maximum of file histories. We can also get information on the limiting profile, i.e. on how much time is spent at each level k (k fixed, n going to infinity). Our investigations are based on the assumption that the size $Y(\lfloor nt \rfloor)$ of the list at time nt (with $0 \le t \le 2$) satisfies a weak convergence property:

$$\frac{Y(\lfloor nt \rfloor) - ny(t)}{\sqrt{n}} \Rightarrow X(t),$$

where y is a symmetric function around 1 and X is a Gaussian process with mean 0 and known covariance C. In all applications to either classical or Knuth-type file histories, that assumption is true.

Let k, the level under study, be fixed, and $\bar{t}$ be the time such that $y(\bar{t}) = k/n$, with $\bar{t} < 1$.

If we consider the time u where Y first hits level k, the density of u is given by

$$\begin{aligned} g(u)du &= \Pr\{\min\{u', Y([nu']) = k\} \in du\} \\ &= \Pr\{\min\{u', X(u') = \sqrt{n}(\frac{k}{n} - y(u'))\} \in du\}, \end{aligned}$$

and so

$$u - \bar{t} = O(\frac{1}{\sqrt{n}}).$$

Thus we examine the behaviour of Y near $\bar{t}$. Locally y can be approximated by a straight line, so that, $y'(\bar{t})$ denoting the slope of y at $\bar{t}$, we have

$$g(u)du \sim \Pr\{\min\{u', X(u') = \sqrt{n}(\bar{t} - u')y'(\bar{t})\} \in du\}.$$

But it is known that locally X behaves like a Brownian Motion [12]. Using classical results on the crossing time of a Brownian Motion and a straight line Cox and Miller [8] (p.221) and an asymptotic analysis, we find that the density of $\tau := \frac{(u-\bar{t})}{\sqrt{n}}$ is just

$$\frac{y'(\bar{t})}{\sqrt{2\pi C(\bar{t},\bar{t})}} e^{-(\tau y'(\bar{t}))^2/2C(\bar{t},\bar{t})} d\tau,$$

where C is the covariance of X. Thus u is a Gaussian variable centered at $\bar{t}$.

We now study the *total time spent* at level k. Let $p(u), q(u), r(u)$ be the probabilities that the next move on the list is an insertion, deletion and query respectively, if we start at $Y([nu])$ at time u. The random walk describing the evolution of the data structure is transient, so that between the time when the size of the structure first hits k, and the time when it last leaves level k, we may consider that $p(u), q(u)$ and $r(u)$ are locally constant and equal to the probabilities p, q and r of insertion, deletion and query at level k.

If the first step leaving level k is a deletion, we are sure of coming back to level k (since $\bar{t} < 1$). If it is an insertion, then the probability of ever returning to level k is q/p.

If we look at the whole file history after hitting level k, the history spends l steps at level k, i steps inserting from k to $k+1$, d steps deleting from k to $k-1$, plus various other operations at other levels. Thus, if $F(z, w, v)$ is the joined generating function

$$F(z, w, v) = \sum_{i,d,l} \Pr\{l, i, d\} z^l w^i v^d,$$

we have when $n \longrightarrow \infty$:

$$F(z, w, v) = qzvF + pw[\frac{q}{p}zF + (1 - \frac{q}{p})] + rzF,$$

which leads to

$$F(z, w, v) = \frac{(p-q)w}{1 - qz(v+w) - rz}.$$

¿From that we can derive the distribution $F(z, 1, 1)$ of the time spent at level k, and similarly of the number of insertions and of deletions from level k. All turn out to be geometric random variables.

5.1 Example: classical priority queues

For priority queues, we have for the average size at t:

$$y(t) = \frac{1}{2}t(2-t),$$

and the covariance of X is given by $C(s,t) = s^2(2-t)^2/4$. Thus $\bar{t} = 1 - \sqrt{1 - 2k/n}$, and $y'(\bar{t}) = \sqrt{1 - 2k/n}$. Adapting now the proof of Lemma 13 in [33], we can prove that

$$p = 1 - \frac{\bar{t}}{2} + O(\frac{1}{n}), \quad q = \frac{\bar{t}}{2} + O(\frac{1}{n}), \quad r = 0.$$

Then we find that

$$E(l) = \frac{1-\sqrt{1-2\frac{k}{n}}}{\sqrt{1-2\frac{k}{n}}} \text{ and } E(i) = E(d) = \frac{1}{\sqrt{1-2\frac{k}{n}}}.$$

5.2 Example: dictionaries in Knuth's model

Similarly we find:

$$y(t) = \left[\sqrt{\frac{3}{2}}(2\cos(\theta/3)-1)/2\right]^{1/2}$$

with $\theta = \arccos[2t(2-t)-1]$. Also $\overline{y'} := y'(\bar{t}) = \sqrt{1-2(\frac{2}{3})^{1/2}(k/n)^{1/2}}$. From [33], we obtain

$$\begin{aligned} p &= \frac{1}{4}(1+2\bar{y}'+\bar{y}'^2) \\ r &= \frac{1}{2}(1-\bar{y}'^2) \\ q &= \frac{1}{4}(1-2\bar{y}'+\bar{y}'^2) \end{aligned}$$

Finally this gives

$$E(l) = \frac{2}{\bar{y}'} \text{ and } E(i) = E(d) = \frac{1+\bar{y}'^2}{2\bar{y}'}$$

6 Hashing with lazy deletion

Hashing with lazy deletion was introduced in [44] as a data structure suitable for example for line-sweep algorithms on a set of segments of the plane. When a new item arrives, (i.e. the line reaches the extremity of a new segment), it is inserted in a random bucket of a separate hash table, along with the x-coordinate of its right extremity. When an item "dies", i.e. the line goes past the right extremity of the segment, it is not removed from the data structure right away, but only at the time of a later insertion within the same bucket. This method presents the advantage of minimizing the number of accesses to the hash table, at the cost of some extra space used. In [44], several models of distribution are suggested to analyze the extra space: two non-stationary models (which are quite similar to one another, so that we will only study the first one, knowing that the same approach works also for the othermodel) and one stationary model.

6.1 The first non-stationary model

In the first non-stationary model, there are n segments in $[0...1]$ drawn independently from the following distribution: segment $s = [\min(x,y), Max(x,y)]$, where x and y are uniform independent r.v.'s in $[0...1]$. Thus the arrival time density is $2n(1-u)du$ and the lifetime z, conditional on u, has density $ndz/(1-u)$. To remove the conditionality, it is convenient to change the time scale. Let the new time t be given by $1-u = e^{-2t}$. It is easy to see that the arrival times now have density $2e^{-2t}dt$, and the service times have unconditional density $e^{-t}dt$.

Two parameters are of interest here: $Need(t)$, the number of items alive at time t (which have to be present in the data structure), and $Use(t)$, the number of items which are actually present in the table at time t. Thus $Waste(t) = Use(t) - Need(t)$ counts the items which are dead but not yet deleted at time t.

6.1.1 Study of $Max_t Need(t)$.

We note that $Need(t)$ depends on n but not on H. Our calculations are asymptotic, when n, the total number of items, goes to infinity. From Louchard [35], Theorem 1, we find that as n grows, $Need(t)$ converges to a Gaussian process:

$$\frac{Need(t) - nz_{Need}(t)}{\sqrt{n}} \Rightarrow Q(t),$$

where $z_{Need(t)}$, the probability that a given item is alive at time t, is defined by

$$z_{Need}(t) = \int_0^t 2e^{-2u}e^{-(t-u)}du = 2e^{-t}(1-e^{-t}).$$

$Q(t)$ is a Gaussian process with mean 0 and covariance

$$C(s,t) = 2e^{-t}(1-e^{-s}) - 4e^{-t}(1-e^{-t})e^{-s}(1-e^{-s})$$

if $s \le t$. It is easy to check that $z_{Need}(t)$ is maximized at $\bar{t} = \ln 2$, where its value is $1/2$. Thus we can rewrite the convergence as

$$Need(t) \sim \sqrt{n}[\frac{\sqrt{n}}{2} + Q(t) + \sqrt{n}(z_{Need}(t) - \frac{1}{2})],$$

and, applying Daniels' theorem, we obtain after some algebra:

Theorem 11

$$Max_t Need(t) \sim \frac{n}{2} + \sqrt{n}M + O(n^{1/6}),$$

where M is characterized by the theorem in section 3.1, with $A = 1$, $B = 1$, $C(\bar{t},\bar{t}) = 1/4$, $[\partial_s C]_{\bar{t}} = 1/2$ and $[\partial_t C]_{\bar{t}} = -1/2$.

6.1.2 Study of $Max_t Use(t)$.

$Use(t)$ does depend heavily on the number of buckets. The larger H is compared to n, the more discrepancy there is between $Max_t Need(t)$ and $Max_t Use(t)$. We assume that the number H of buckets also goes to infinity, and there are two cases: either $H = rn$, with r fixed, or $H = \alpha(n)n$, with $\alpha(n) \to 0$.

We must now study $M = Max_t Use(t)$. Let $Waste(k)$ be the time during which customer k was dead but not yet deleted. The distribution of $Waste(k)$ is known:

$$\Pr\{Waste_k > t - y\} = \left[1 - \frac{e^{-2t} - e^{-2y}}{H}\right]^{n-1}.$$

If $H = rn$, we use the proof technique presented in [35] to get a weak convergence of

$$\frac{Use(t) - nz(t)}{\sqrt{n}} \Rightarrow Q(t)$$

to some process $Q(t)$, where $nz(t)$ is the average number of items in use at time t: $z(t)$ is the probability that a fized item is in Use at time t. The item is alive at time t if it arrived at $u < t$ and lived for more than $t - u$; it is dead but not deleted at time t if it arrived at time u, died at time $y < t$, and was in waste for longer than $t - y$. Thus, approximating $Waste_k$, we get

$$z(t) = \int_0^t 2e^{-2u}e^{-(t-u)}du + \int_0^t 2e^{-2u}du \int_u^t e^{-(y-u)}e^{-(e^{-2y}-e^{-2t})/r}dy.$$

The limit process $Q(t)$ has mean 0, and we must calculate its covariance.

Let $\eta_i(s)$ be the random variable which is 1 if customer i is in use at time s, and 0 otherwise. Then the number $Use(s)$ of customers in use at time s satisfies

$$Use(s) = \sum_{1 \le i \le n} \eta_i(s).$$

Thus the covariance is given by

$$\mathrm{Cov}(Use(s)Use(t)) = \sum_{1 \le i \le n} E(\eta_i(s)\eta_i(t)) + 2\sum_{i<j} E(\eta_i(s)\eta_j(t)) - n^2 z(s)z(t)$$

for $s \le t$. The first part of the sum can be expressed easily as a sum of integrals. The second term depends on whether i and j are in the same bucket or not (which events have probabilities $1/H$ and

$1 - 1/H$). After carefully developing in $1/n$, we obtain

$$
\begin{aligned}
\mathrm{Cov}(Use(s)Use(t)) = n\,[& \\
& 2e^{-t}(1-e^{-s}) + \int_0^s 2e^{-2u}du \int_u^t e^{-(y-u)} e^{-\frac{e^{-2y}-e^{-2t}}{r}} dy \\
+ \; & \eta(s)z(t) + \eta(t)z(s) + \eta(s)\eta(t) \\
+ \; & \frac{1}{r}\int_0^s 2e^{-2u}du \int_0^t 2e^{-2v}dv \int_u^s e^{-(x-u)}dx \int_v^t e^{-(y-v)}dy e^{-A([xs]\cup[yt])/r} \\
- \; & \frac{1}{r}\int_0^s 2e^{-2u}du \int_u^s e^{-(y-u)} e^{-(e^{-2y}-e^{-2s})/r} dy \int_0^t 2e^{-2u}du \int_u^t e^{-(y-u)} e^{-(e^{-2y}-e^{-2t})/r} dy \\
- \; & z(s)z(t)] + O(1),
\end{aligned}
$$

with: $z(t)$ given as above, $\eta(s)$ defined by

$$\eta(s) = \int_0^s 2e^{-2u}du \int_u^s e^{-(x-u)} e^{-(e^{-2x}-e^{-2s})/r} \frac{e^{-2x}-e^{-2s}}{r} dx,$$

and $A([xs]\cup[yt])$ is the measure of the union of intervals $[xs]$ and $[yt]$ in the distribution $A(x) = 1 - e^{-2x}$.

To apply Daniels' theorem, we must find $\bar{t}$ such that $z'(\bar{t}) = 0$. From [44], it is known that $\bar{t}$ only exists if $r < 0.84$.

If $r < 0.84$, we can apply Daniels' theorem.

If $r \geq 0.84$, then $z(t)$ is an increasing function, and we know from Iglehart [24] that $Max_{s\leq t}Use(s)$ is asymptotically distributed as $Use(t)$. Letting t go to infinity gives an equivalent of the Maximum.

Finally we have:

Theorem 12 *Assume that $H = rn$.*

1. *If $r \geq 0.84$, then*

$$Max_t Use(t) \sim Cn + \sqrt{n}\sqrt{C(1-C) + C'}X,$$

where $X = \mathcal{N}(0,1)$ is the classical Gaussian random variable, C is defined to be $2\int_0^\infty e^{-x}(1-e^{-x})e^{-e^{-2x}/r}dx$, and C' (which can be evaluated numerically) is given by:

$$
\begin{aligned}
C' = \; & 2\eta(\infty)z(\infty) + \eta^2(\infty) \\
+ \; & \frac{1}{r}\int_0^\infty 2e^{-2u}du \int_0^\infty 2e^{-2v}dv \int_u^\infty e^{-(x-u)}dx \int_v^\infty e^{-(y-v)}dy e^{-A([xs]\cup[yt])/r} \\
- \; & \frac{1}{r}\left[2\int_0^\infty e^{-x}(1-e^{-x})e^{-e^{-2x}/r}dx\right]^2.
\end{aligned}
$$

2. *If* $r < 0.84$, *then*

$$Max_tUse(t) \sim n[2e^{-\bar{t}}(1-e^{-\bar{t}})+r] + \sqrt{n}M + O(n^{-1/6}),$$

where $\bar{t}$ *is the solution of*

$$2\int_0^{\bar{t}} e^{-x}(1-e^{-x})e^{-(e^{-2x}-e^{-2\bar{t}})/r}dx = r,$$

and M *is given by Daniels' theorem, with* $B = [4e^{-3\bar{t}}(1-2e^{-\bar{t}})/r]^{-1/3}$, *and* $C(\bar{t},\bar{t})$, $[\partial_s C]_{\bar{t}}$ *and* $[\partial_t C]_{\bar{t}}$ *can be written in terms of complicated multiple integrals.*

In the case where $H = \alpha(n)n$, with $\alpha(n) \to 0$, the arrival times in a bucket behave, in first approximation, like a Poisson process. Our goal here is again to apply Daniels' theorem. At each step of the calculation, we write the quantities as power series in α, so that we can neglect high powers of α, and the expressions remain simple enough. Apart from some technical complications, the approach is the same as for the case where $H = rn$. Our calculations are accurate if $\alpha(n) = o(1/\sqrt{n})$. If $\alpha(n)$ decreased more slowly than that, it would still be possible to get some results, but that would require writing the power series more accurately.

Theorem 13 *Let* $H = \alpha(n)n$, *with* $\alpha(n) = o(1/\sqrt{n})$. *Then we have*

$$Max_tUse(t) \sim n(\frac{1}{2}+\alpha-4\alpha^2+O(\alpha^3)) + \sqrt{n}M + O(n^{-1/6}),$$

where M, *given by Daniels' theorem, is asymptotically Gaussian, with mean*

$$E(M) = 0.99615n^{-1/6}2^{2/3} + O(n^{-1/3}).$$

We have $A = 2 + O(\alpha)$ *and* $B = 1 + O(\alpha)$,

$$C(\bar{t},\bar{t}) = \frac{1}{4} + 2\alpha + O(\alpha^2)$$

$$[\partial_s C]_{\bar{t}} = 1 + O(\alpha)$$

$$[\partial_t C]_{\bar{t}} = -1 + O(\alpha).$$

6.2 The second non-stationary model

Van Wyk and Vitter suggested another non-stationary model, in which arrivals and deaths are not symmetric. Here again n segments in $[0\ldots 1]$ are drawn independently from a given distribution: segment $[x\ldots y]$ is constructed by drawing x uniformly in $[0\ldots 1]$, and then drawing y uniformly in $[x\ldots 1]$. Up to a change of scale, we can assume that the densities of the arrival and service times are $e^{-t}dt$. From then on, the proof mirrors that of the previous subsection; the only difference in the calculations is that the arrival times have density $e^{-t}dt$ instead of $2e^{-2t}dt$, so that there is no critical value like 0.84: there is always a point at which $z(t)$ is maximized.

Theorem 14

$$Max_t Need(t) \sim \frac{n}{e} + \sqrt{n}M + O(n^{1/6}),$$

where M is characterized by Daniels' theorem, with $A = 2/e$, $B = (1/e)^{-1/3}$, $C(\bar{t},\bar{t}) = (1-1/e)/e$, $[\partial_s C]_{\bar{t}} = 1/e$ and $[\partial_t C]_{\bar{t}} = -1/e$.

Theorem 15 *If $H = \alpha(n)n$, with $\alpha(n) = o(1/\sqrt{n})$, then we have:*

$$Max_t Use(t) \sim n(\frac{1}{e} + \alpha + O(\alpha^2)) + \sqrt{n}M + O(n^{1/6}),$$

where M is characterized by Daniels' theorem, with $A = 4/e + O(\alpha)$, $B = (1/e)^{-1/3} + O(\alpha)$,

$$C(\bar{t},\bar{t}) = \frac{1}{e}(1-\frac{1}{e}) + \alpha(1+\frac{2}{e}) + O(\alpha^2),$$

$[\partial_s C]_{\bar{t}} = 2/e + O(\alpha)$, and $[\partial_t C]_{\bar{t}} = -2/e + O(\alpha)$.

7 Conclusion

Diffusion techniques allowed us to derive several new results on data structures maxima. Many problems remain open, and are the object of work in progress. Let us mention the symbol table, (the probabilistic properties of which are unknown), the $G/G/1$ detailed maximum properties, the $G/G/\infty$ case, with $\ln(\lambda t) = \Omega(\frac{\lambda}{n})$. Let us remark that, in a recent report [2], Aldous, Hofri and Szpankowsky have analysed Hashing with lazy deletion in the stationary case and proved open conjectures introduced in [37] and [44].

Acknowledgments

This work has been supported in part by the P.R.C. "Mathématiques et Informatique".

References

[1] ABRAMOWITZ, M. and STEGUN, I.A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*; 1965, Dover

[2] ALDOUS, D., HOFRI, M. and SZPANKOWSKI, W. *Maximum Size of a Dynamic Data Structure: Hashing with Lazy Deletion Revisited*; Purdue University, CSD-TR-992, July 1990

[3] ANDERSON, C.W. *Extreme Value Theory for a class of Discrete Distributions with Applications to some Stochastic Processes*; J. Appl. Prob., 1970, **7**, 99-113

[4] BERMAN, S.M. *Sojourns and Extremes of Stationary Processes*; The Annals of Probability, 1982, **10**,1, 1-46.

[5] BERMAN, S.M. *Sojourns and Extremes of a Diffusion Process on a Fixed Interval*; Adv. Appl. Prob., 1982, **14**, 811-832.

[6] BILLINSLEY, P. *Convergence of Probability Measures*, 1968 Wiley

[7] COHEN, J.W. *The Single Server Queue*, Revised Edition, 1982, North-Holland

[8] COX, D.R. and MILLER, H.D. *The Theory of Stochastic Processes*; 1965, Chapman and Hall

[9] CSAKI, E., FOLDES, A. and SALMINEN, P. *On the Joint Distribution of the Maximum and its Location for a Linear Diffusion*; Ann. Inst. Henri Poincare, 1987, **23**, 2, 179-194

[10] DANIELS, H.E. and SKYRME, T.H.R. *The Maximum of a Random Walk whose Mean Path has a Maximum*; Adv. Appl. Prob., 1985, bf 17, 85-99

[11] DANIELS, H.E. *The Maximum of a Gaussian Process whose mean Path has a Maximum, with an Application to the Strenfth of Bundles of Fibres*; Adv. Appl. Prob., 1989, **21**, 315-333

[12] DURBIN, J. *The First-Passage Density of a Continuous Gaussian Process to a General Boundary*; J. Appl. Prob., 1985, **22**, 99-122

[13] FELLER, W. *Introduction to Probability Theory and its Applications - Vol.II*; Wiley, 1971

[14] FLAJOLET, P., FRANÇON, J. and VUILLEMIN, J. *Sequence of Operations Analysis for Dynamic Data Structures*; Journal of Algorithms, 1(2) (1980).

[15] FLAJOLET, P. and PRODINGER, H. *Register Allocation for Unary-Binary Tress*, SIAM J. Comput.

[16] FLAJOLET, P. and STEYAERT, J.M. *A Branching Process Arising in Dynamic Hashing, Trie Searching and Polynomial Factorization*; 9th Int. Coll., Aut. Lang. Prog. Aarhus 1982. Lecture notes in Comp. Sc. **140**, 239-251

[17] FLAJOLET, P. *Approximate Counting: a Detailed Analysis*; BIT, 1985, **25**, 113-134

[18] FLAJOLET, P., PUECH, C. and VUILLEMIN, J. *The Analysis of Simple List Structures*, Inform. Sci., 1986, **38**, 121-146

[19] FRANÇON, J. *Combinatoire des structures de données*, Thèse de doctorat d'Etat, Université de Strasbourg, 1979

[20] FRANÇON, J., RANDRIANARIMANANA, B. and SCHOTT, R. *Analysis of Dynamic Algorithms in Knuth's Model*; Theoretical Computer Science, 1990, **72**, 147-167

[21] IGLEHART, D.L. *Limiting Diffusion Approximations for the Many Server Queue and the Repairman Problem*; J. Appl. Prob., 1965, **2**, 429-441

[22] IGLEHART, D.L. *Weak Convergence in Queueing Theory*; Adv. Appl. Prob., 1973, **5**, 570-594

[23] IGLEHART, D.L. *Weak Convergence of Compound Stochastic Process, I*; Stoch. Proc. and their Appl., 1973, **1**, 11-31

[24] IGLEHART, D.L. *Extreme values in the GI/G/1 Queue*; The Annals of Math. Stat.; 1972, **43**, 2, 627-635

[25] JAESCHKE, D. *The Asymptotic Distribution of the Supremum of the Standardized Empirical Distribution Function on Subintervals*; The Annals of Math. Stat., 1979, **7**, 1, 108-115

[26] JOHNSON, N.L.and KOTZ, S. *Distribution in Statistics: Continuous Univariate Distributions*; 1970, Wiley

[27] KEILSON, J. and ROSS, H.F. *Passage Time Distributions for Gaussian Markov (Ornstein-Uhlenbeck) Statistical Processes*; Selected Tables in Mathematical Statistics, **III**, 1975

[28] KENYON-MATHIEU, C.M. and VITTER, J.S. *General methods for the analysis of the maximum size of dynamic data structures*; Proceed. ICALP'89, Lect. Notes in Comp. Sc. **372**, 473-487

[29] KNUTH, D.E. *Deletions that Preserve Randomness*; Trans. Software Eng., 1977, 351-359

[30] LOUCHARD, G. *Brownian Motion and Algorithm Complexity*; BIT, 1986, **26**, 17-34

[31] LOUCHARD, G. *Exact and Asymptotic Distributions in Digital and Binary Search Trees*; Theor. Inf. and Appl., 1987, **21**, 479-496

[32] LOUCHARD, G. *Random Walks, Gaussian Processes and List Structures*; Theor. Comp. Sc., 1987, **53**, 99-124

[33] LOUCHARD, G., SCHOTT, R. and RANDRIANARIMANANA, B. *Dynamic Algorithms in D.E. Knuth's Model : a Probabilistic Analysis*; Proceed. ICALP'89, Lect. Notes in Comp. Sc. **372**, 521-533
(full paper to appear in Theor. Comp. Sc.)

[34] LOUCHARD, G. *Kac's Formula, Levy's Local Time and Brownian Excursion*; J. Appl. Prob., 1984, **21**, 479-499

[35] LOUCHARD, G. *Large Finite Population Queueing Systems. Part I: the Infinite Server Model*; Commun. Statist.-Stoch. Models, 1988, **4**, 3, 473-505

[36] LOUCHARD, G. and SCHOTT, R. *Probabilistic Analysis of Some Distributed Algorithms*; Proceed. CAAP'90, Lect. Notes in Comp. Sc. **431**, 239-253
(full paper to appear in Random Structures and Algorithms)

[37] MATHIEU, Cl. *Comparaison de modèles combinatoires et probabilistes: deux exemples en Analyse d'Algorithmes*; Thèse de doctorat, Univ. Paris-Sud, Centre d'Orsay, 1988

[38] MORRISON, J., SHEPP, L.A., and VAN WYK, C.J. *A queuing analysis of hashing with lazy deletion*; SIAM J. on Comp. 16 (6) (1987).

[39] RANDRIANARIMANANA, B. *Complexité des structures de données dynamiques*; Thèse de doctorat, Université de Nancy 1 , 1989

[40] SADOWSKY, J.S. and SZPANKOWSKI, W. *The Probability of Large Queue Lenths and Waiting Times in a Heterogeneous GI/GI/c Queue*; Preprint

[41] SALMINEN, P. *On the First Hitting Time and the Last Exit Time for a Brownian Motion to/from a Moving Boundary*; Adv. Appl. Prob. , 1988, **20**, 411-426

[42] SERFOZO, R.F. *Extreme Values of Birth and Death Processes and Queues*; Stoch. Proc. and their Appl., 1988, **27**, 291-306

[43] SERFOZO, R.F. *Extreme Values of Queue Length in M/G/1 and G/M/1 System*; Math. Op. Res., 1988, **13**, 349-357

[44] VAN WYK, C.J. and VITTER, J.S. *The Complexity of Hashing with Lazy Deletion*; Algorithmica, 1986, **1**, 17-29

A A direct computation of Daniels' formula

Daniel's results are based on the following hitting time density for a B.M. X(t).
Let $w(t) = m + \sqrt{n}\, f(t)$, with $f(t_o) = f'(t_o) = 0$. Daniels has obtained the density (see Daniels and Skyrme [10]):

$$g(t)dt = Pr[\min(t' : X(t') = w(t')) \in dt] = \frac{e^{-\frac{[w(t)]^2}{2t}}}{\sqrt{2\pi t}} \mu(t)dt \tag{1}$$

where

$$\mu(t) = n^{-1/3}(2\beta)^{1/3} F\left\{n^{1/3}[2\beta]^{2/3}(\bar{t} - t)\right\}\left(1 + 0(n^{-1/3})\right)$$

with F given by (25), $\beta := f''(\bar{t})/2$ and $\bar{t}$ is such that $h'(\bar{t}) = h(\bar{t})/\bar{t}$.

(i) We will firstly extract all dominant terms from (1). As in [10], set

$$x = n^{1/3}(2\beta)^{2/3}(t_o - t) = 0(1)$$

They have computed

$$\bar{t} - t_o = n^{-1/2}\frac{m}{t_o 2\beta} + 0(n^{-2/3})$$

so

$$\bar{t} - t = \frac{xn^{-1/3}}{(2\beta)^{2/3}} + \frac{n^{-1/2}m}{2\beta t_o} + 0(n^{-2/3})$$

Expanding $G(x)$ as given in Sec.4.1, we obtain

$$G(x) = \frac{1}{2^{2/3}}\sum_{k=o}^{\infty}\frac{e^{-\lambda_k x/2^{1/3}}}{A_i'(\lambda_k)}$$

where λ_k are the zeros of $A_i(x)$. A detailed expansion of (1) now leads to

$$\begin{aligned} g(t) &= \frac{(2\beta)^{1/3}}{\sqrt{2\pi t_o}\, 2^{2/3}} \exp\left[\frac{-m^2}{2t_o} + \frac{x^3}{6}\right] \\ & \sum_{k=o}^{\infty} \frac{\exp\left[\frac{-\lambda_k}{2^{1/3}}\left(x + \frac{m}{t_o(2\beta)^{1/3}n^{1/6}}\right)\right]}{A_i'(\lambda_k)}(1 + 0(n^{-1/3})) \end{aligned} \tag{2}$$

(ii) Such a simple formula as (2) should be explained in direct hitting time density for B.M. . This will be done with a technique introduced by Salminen [41]. Let us first remark that, to derive

(2), it is enough to limit ourselves to a second-order boundary (error is within $O(n^{-1/3})$). So we can simply use

$$w(t) = m + \sqrt{n}\beta(t - t_o)^2$$

with $\bar{x} = w(0) = m + \beta\sqrt{n}\, t_o^2$. Our hitting problem can now be transformed into

$$g(t)dt = Pr_{\bar{x}}[\min(t' : X(t') = h(t')) \in dt]$$

with $h(t) = -\sqrt{n}\,\beta(t - t_o)^2 + \sqrt{n}\,\beta t_o^2$ such that $h(0) = 0$.
This also gives $h'(t) = -2\sqrt{n}\,\beta(t - t_o)$; $h'' = -2\beta\sqrt{n}$.
By Salminen [41] (2.6) and (3.9), we now obtain

$$\begin{aligned} g(t) &= 2\left(\frac{\beta\sqrt{n}}{2}\right)^{2/3} \exp\left[h'(0)\bar{x} - \frac{1}{2}\int_o^t [h'(s)]^2 ds\right] \\ &\quad \sum_{k=o}^{\infty} e^{\lambda_k (2\beta^2 n)^{1/3} t} \frac{A_i\left[\lambda_k + 2(\frac{\beta\sqrt{n}}{2})^{1/3}\bar{x}\right]}{A'(\lambda_k)} \end{aligned} \tag{3}$$

But $h'(0) = 2\sqrt{n}\,\beta t_o$ and

$$\begin{aligned} -\frac{1}{2}\int_o^t [h'(s)]^2 ds &= -\int_o^t 2n\beta^2(s - t_o)^2 ds \\ &= -2n\beta^2\left[\frac{(t - t_o)^3}{3} + \frac{t_o^3}{3}\right] = \frac{x^3}{6} - \frac{2n\beta^2 t_o^3}{3} \end{aligned}$$

By Abramowitz and Stegun [1] (10.4.59) we know that, for large z,

$$A_i(z) \sim \frac{1}{2}\pi^{-1/2} z^{-1/4} e^{-\xi}(1 + 0(\frac{1}{\xi}))$$

with $\xi := \frac{2}{3}z^{3/2}$.
From (3), we must use

$$z = \lambda_k + 2^{2/3} n^{2/3} \beta^{4/3} t_o^2 + 2^{2/3}\beta^{1/3} m n^{1/6}$$

so that

$$\begin{aligned} \xi &= \frac{4}{3}n\beta^2 t_o^3\left[1 + \frac{3}{2}\frac{m}{\beta t_o^2\sqrt{n}} + \frac{3}{8}\frac{m^2}{\beta^2 t_o^4 n} + \frac{3}{2}\frac{\lambda_k}{2^{2/3} n^{2/3}\beta^{4/3} t_o^2}\right. \\ &+ \left.\frac{3}{4}\frac{m\lambda_k}{\beta^{7/3} t_o^4 2^{2/3} n^{7/6}}\right]\left(1 + 0(\frac{1}{n^{1/3}})\right) \end{aligned}$$

Also, we deduce

$$\lambda_k(2\beta^2 n)^{1/3} t = \lambda_k(2\beta^2 n)^{1/3} t_o - \frac{\lambda_k x}{2^{1/3}}$$

$$h'(0)\bar{x} = 2\beta\sqrt{n}\, t_o \bar{x} = 2\beta\sqrt{n}\, m t_o + 2n\beta^2 t_o^3$$

Identification of (3) with (2) is now routine.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique

ISSN 0249-6399