Balls and buckets for discrete time Markov chain generation †

Héctor Cancela Bosi
cancela@fing.edu.uy
Investigación Operativa - InCo - Pedeciba Informática
Facultad de Ingeniería, Universidad de la República
Montevideo, Uruguay

Abstract

This work presents an extension of the balls and buckets description formalism to the case of discrete time systems. This formalism allows for flexible modeling and compact representation of Markovian systems, supporting automatic generation of both the state space and the transition matrix of the underlying canonical Markov chain. A discrete time balls and buckets interpreter has been implemented in the BB library, which provides different basic services (Markov chain generation, storage, exact evaluation, and simulation) supporting system description and performance and dependability quantitative analysis.

As an application example, we present the balls and buckets model and the performance analysis of an ATM switch architecture, which is of interest in high performance network design. By using simulation techniques, we reproduce results (already established by other methods in the literature) showing that, given a fixed load, the performance of this kind of switches is highly sensitive to the characteristics of the data packets arrival process.

Keywords: system modeling, simulation, performance and dependability evaluation, distributed systems and computer networks.

1 Introduction

The complexity of the new generations of computer and communications systems and the fact that many activities depend on their good performance has promoted a growing attention to modeling and evaluation techniques applicable to these systems. The main properties are the performance and the dependability, this last one being of particular interest in so-called critical systems, for example systems controlling nuclear generators, commercial, military or spatial flight systems, etc. Other interesting measures include the reliability, the availability, the performability (performance-dependability combined measure), etc. To study these measures we model the systems using stochastic processes, typically in a Markovian setting. Many numerical techniques have been developed to evaluate Markovian processes. The main problem is that the state space size of the process grows

†This work was supported by grants from Action U93E03 of ECOS - Scientific Cooperation between France and Uruguay Program, and from CONICYT (Uruguay)
(often very fast) with the complexity and detail level of the model, which makes it impossible for the user to directly specify the corresponding process. This job requires a systematic modeling approach, with the use of high level interfaces offering a compact representation of the studied system.

Within the Markovian setting, we find two main model families, which correspond to the use of a continuous or discrete time parameter. In the continuous time setting, there are a number of tools based on formalisms which allow to give a compact representation of the system studied, and generate afterwards the underlying Markov process. We can find, among others, tools based on Petri nets [3, 10, 11, 13, 19], stochastic activity networks [17, 19, 20], queuing systems [16], balls and buckets [8, 21], Markovian event systems [14], automata networks [18], process algebras [15], logical frameworks [9], hybrid logical-object oriented approaches [4, 6], and specific "machine-repairmen"-like languages as SAVE [5].

The discrete time setting has been less explored, but has also been getting more attention recently, in response to recent changes in the communication networks technology. In particular, it is necessary to employ a discrete time parameter in order to model more precisely both software and hardware behavior. A case of particular interest is the modeling of high performance communication networks based on ATM. This need has motivated recent works presenting discrete time formalisms (discrete time automata networks [18], discrete time stochastic Petri nets [12], discrete time queuing networks [24]), which can be used to study the behavior of those systems. Nevertheless, most existing modeling tools restrict themselves to only continuous time formalisms, and do not include discrete time models.

In this work we present an extension of the balls and buckets formalism (previously employed for continuous time system modeling) to the discrete time case. This formalism can be used to give system descriptions more compact than those based on Markov chains, while maintaining the same expressive power (the classes of systems that can be represented are equivalent). To extend the formalism to the discrete time case, it is necessary to expand its semantics, to take into account the fact that many changes of the system state can take place simultaneously, at the same time unit. This extension has been incorporated into the BB model generation library [7, 8], which previously only supported continuous time models. The library is composed by a set of classes and methods for balls and buckets model construction and evaluation. In particular, it is possible to find the Markov chain equivalent to any specific BB model, opening access to the battery of existing analytical methods; also it is possible to use simulation methods, when the size of the model makes impractical (or impossible) the first alternative. As an example, we have taken a discrete time model of resource contention for a specific ATM switch architecture.

This paper is organized as follows. Next section presents the discrete time balls and buckets formalism. In Section 3, we discuss the implementation of this formalism in the BB library. In Section 4 we develop as an application example the modeling and performance evaluation of an ATM switch. Finally, some conclusions and future work perspectives are given in Section 5.

2 Discrete time balls and buckets formalism

2.1 Informal presentation and example

The balls and buckets modeling formalism has been proposed by Stewart [21] as a compact means for representing continuous time Markovian systems by the use of a reduced set of simple rules, allowing for flexible modeling and automatic state space and infinitesimal generator matrix construction.
In its original form it is not suitable to discrete time system modeling. In order to cope with this case, we have developed an extended version of the formalism. While preserving the main philosophy of the previous approach, it has been necessary to modify and widen the formalism semantics, specially because it becomes necessary to explicitly take into account the possibility that many system state modifications may take place in a single discrete time unit. The differences between the continuous and the discrete time versions are enough to preclude an easy unified presentation, so from now on we will refer exclusively to the new discrete time formalism.

In the discrete time formalism we propose, the system state is represented by a set of \( L \) buckets, each containing a number of balls \( n[i] \) (with \( n[i] \in \{0, \ldots, nmax[i]\}, i = 1, 2, \ldots, L \)). The balls can move among the different buckets. The movement of a ball from a bucket \( i \) to a bucket \( j \) is called a transfer from source \( i \) to destination \( j \), and can happen only if \((i, j) \in L \) (the transfer belongs to the list of admissible transfers), and also \( n[i] > 0 \) and \( n[j] < nmax[j] \), or \( n[j] = nmax[j] \) if \( i = j \) (enabling condition). As a result of a transfer from \( i \) to \( j \), the number of balls \( n[j] \) goes up by 1 unit and \( n[i] \) goes down also by 1. In a given time unit, a bucket may be the source of at most one transfer, but it may be the destination of many. The transfers happen with given probabilities, which may depend upon the global system state. The sum of the probabilities of transfers having the same source bucket must be less or equal to one. A transfer may be followed by an arbitrary modification of the buckets contents, which takes place instantaneously. Since at the same time many transfers (having different source buckets) may happen, we consider that these transfers take place ordered by their input bucket position, transfers having lower numbered input buckets taking place before higher numbered buckets (this is the same as giving arbitrary priorities to the buckets, and re-ordering them). In this case a transfer is possible if the enabling condition was true at the cycle start and if it is still true after consideration of the effects of all transfers having greater priority (i.e., having smaller input bucket number).

A transition is a system state change, resulting from the composition of all transfers and instantaneous modifications which take place in a single time unit.

![Figure 1: Geo/Geo/1/K queuing system](image)

As an example let us consider a Geo/Geo/1/K discrete time queuing system, which is shown in Figure 1. In each time unit a new customer can arrive with probability \( p \) (the arrival law is a Bernoulli process with parameter \( p \)). The service time is a random variable with geometric distribution and parameter \( r \) (we will denote \( \bar{p} = 1 - p \), and \( \bar{r} = 1 - r \)). The queue has a fixed capacity of \( K \) customers. We suppose that in this system “arrivals see departures”, that is to say during the same time unit departures from the system are processed before arrivals.

The following parameters define (informally) a balls and buckets model for this system:

- \( L = 2 \): two buckets, the first one corresponding to the customers in the queue (including the one being serviced, if any) and the second one corresponding to the system environment;
• \( n_{\text{max}}[1] = K + 1; \)
  \( n_{\text{max}}[2] = K + 1; \)
• \( \mathcal{L} = \{(2, 1), (1, 2)\} \): the admissible transfers;
• transfer modeling a customer arrival: from bucket 2 to bucket 1, with probability \( p; \)
• transfer modeling the end of a service at the station:
  from bucket 1 to bucket 2, with probability \( r; \)
• initial state: \( (0, K + 1); \)
• no instantaneous modifications.

In Figure 2, we can see the Markov chain which corresponds to this system.

![Markov chain diagram](image)

Figure 2: Markov chain corresponding to the Geo/Geo/1/K queue

2.2 Formal Definition

A discrete time balls and buckets model is defined by the 6-uple \( (L, n_{\text{max}}, \epsilon_0, \mathcal{L}, \lambda, \Gamma) \), where:

• \( L \geq 1 \) is the number of buckets;
• \( n_{\text{max}} \in \mathbb{N}^L \) is the maximum number of balls in each bucket;
• \( \epsilon_0 \in \mathbb{N}^L, 0 \leq \epsilon_0[i] \leq n_{\text{max}}[i] \ \forall i \in \{1, \ldots, L\} \), is the initial state;
• \( \mathcal{L} \subseteq L^2 \) is the list of admissible transfers;
• \( \lambda \in (\mathbb{N}^L \rightarrow [0, 1])^\mathcal{L} \) is the transfer probability function (as function of the system state). For each \( \bar{n}, i, j \), such that \( (i, j) \in \mathcal{L}, 0 < n[i] \) and \( n[j] < n_{\text{max}}[j] \) (or \( n[j] = n_{\text{max}}[j] \) if \( i = j \)),

\[
\lambda[i, j](\bar{n}) = \text{ transfer probability of a ball from bucket } i \text{ to bucket } j;
\]

If \( (i, j) \notin \mathcal{L}, n[i] = 0 \) or if \( n[j] = n_{\text{max}}[j] \), by definition \( \lambda[i, j](\bar{n}) = 0 \). Also we always have

\[
\sum_j \lambda[i, j](\bar{n}) \leq 1, \ \forall i, \forall \bar{n}.
\]
\[ \Gamma \in \left( \mathbb{N}^L \to \left( \mathbb{N}^L \times [0,1] \right)^* \right)^L \] is the instantaneous modifications function; for each \( \bar{n} \), resulting from a transfer from \( i \) to \( j \), \( \Gamma \) gives the set of all possible instantaneous modifications:

\[
\Gamma[i,j](\bar{n}) = \begin{cases} 
(\bar{n}, 1) & \text{if there is no instantaneous modification;} \\
(\bar{m}_1, p_1), \ldots, (\bar{m}_H, p_H), (\bar{n}, 1 - \sum_{h=1}^{H} p_h) & \text{where } \bar{m}_1, \ldots, \bar{m}_H \text{ are the accessible states from state } \bar{n} \text{ (after a transfer from bucket } \bar{n} \text{ to bucket } \bar{j} \text{). The probability that the transition to } \bar{m}_h \text{ is chosen is } p_h. 
\end{cases}
\]

It is easy to show that discrete time balls and buckets models are equivalent to homogeneous discrete time Markov chains with finite state space. If we generalize the balls and buckets formalism allowing for the use of buckets with a non-bounded number of balls, we have the class of homogeneous discrete time Markov chains (with state space size possibly being infinite). In this case, it is not possible to generate the state space and the transition matrix, but it is possible to use techniques (such as simulation or bounding methods) that give approximations to useful system measures [1, 2, 22].

\section{Implementation}

\subsection{Architecture of the BB library}

The BB library, presented in [7] and [8], is a set of classes and methods implemented in C++ which, given a system dependability or performance model, build the underlying Markov process which can then be used to compute the desired measures. The library is divided into two main layers. The first layer (which we call the core of the library) is independent of the chosen modeling paradigm. The second layer, which we call the BB interpreter, is a specialization of the core for the treatment of balls and buckets models. These two layers give a number of services, which can be used by a third layer corresponding to user programs. In particular, it is possible to develop translators which can generate a balls and buckets model from a description of a system written in another formalism (as for example GSPNs).

The core of the library is composed by a set of generic classes and methods, which are independent from the modeling formalism. They provide basic abilities such as model generation, display, storage, and simulation.

At this level, a system is described by the function giving all possible transitions from any state, and their associated probabilities. Also the initial state should be known. Using this information, the core can build the set \( \mathcal{S} \) of accessible states and the equivalent Markov process \( X \) with state space \( \mathcal{S} \) and transition matrix \( M \).

The implemented algorithm builds the process transition matrix incrementally, by a BFS search from the initial state. This technique has a number of advantages: for example, only states that can be reached from the initial state will be taken into account in the transition matrix (which is much more efficient than considering the Cartesian product of the system’s individual states, which is the state space used in other modeling tools).

The second layer of the library contains all the necessary information to store and manipulate the states of a system modeled in terms of balls and buckets. This includes the implementation of virtual methods defined at the preceding layer. The user specifies the model as a 6-uple
(L, \bar{m}, \bar{a}, \bar{L}, \lambda, \Gamma). The functions \lambda and \Gamma are given as C++ procedures, which allows a great specification flexibility. The system is represented as a \textit{babSystem} class object, whose methods give transparent access to the library services.

3.2 Transition generation algorithm

The interaction of the two levels of the library is based on an algorithm which from a \textit{balls and buckets} model and a state \( \mathcal{F} \) can give all transitions, with their respective probabilities. This algorithm takes the system description in terms of transfers and instantaneous modifications, and reconstructs its low level behavior.

In the BB library we have implemented such an algorithm, which can be seen in Figure 3, which stores in a list \( \Delta \) the low level information built from the \( \lambda \) and \( \Gamma \) functions specified in the \textit{balls and buckets} model.

To obtain all the combinations of transfers and instantaneous modifications, we consider in order all the buckets, and at each step we apply all the transfers (and corresponding instantaneous modifications) having as source the fixed bucket, over all previously generated states. In this way, we find the new accessible states (with their probability). We multiply the probability of the old states by their probability of not being the source of a transfer in the considered step. For the new step, we add the new states to the set \( \Delta \).

4 ATM Switch Model

4.1 Switch Architecture

The Asynchronous Transfer Mode (ATM) is a technology currently considered for high performance communication networks [23]. These networks must be capable of carrying a large number of highly bursty input links, such as voice, video, and large file transfer. The traffic in these networks is segmented in fixed size packets (or cells). The time is also slotted, each slot being equal to the time necessary to transmit a single cell. An important ATM component is the switch or multiplexer. There are three class of ATM switch architectures: shared-memory, shared-medium, and space-division.

In this section, we will study a shared-medium ATM switch presented in [25], which appears in Figure 4. In this model, all cells arriving from the input links are routed to output links over a common high speed medium, such as a parallel bus. There is some internal storage space (input and output buffers). If an input buffer is full, arriving cells will be lost; if an output buffer is full, there will be blocking at the routing level.

The following are the switch characteristics:

- **Input links**: there are \( N \) input links, each generating cells following an interrupted Bernoulli process (IBP) as follows: each input link has two states, Active and Passive. In a given time slot, a passive input link \( i \) can go to active state with probability \( 1 - q[i] \); and an active link can go to passive state with probability \( 1 - p[i] \). When an input link goes from passive to active state, it generates a cell; while an input link is active, it can generate a cell at each time slot with probability \( \alpha[i] \).

- **Input buffers**: there are \( N \) input buffers of size \( M \), each buffer being associated to an input links. If there is an arrival from input link \( i \) and the corresponding buffer is full, the cell is
Input: state $\vec{e}$
function $\lambda$
function $\Gamma$
Output: list $\Delta$

Begin

$\Delta := \{(\vec{e}, 1)\}$

Repeat for all buckets $i$

$\Delta_i := \emptyset$

Repeat for all pairs $(\vec{e}_k, p_k)$ in $\Delta$

$\Delta_{ik} := \emptyset$

Repeat for all transfers $i \rightarrow j$

enabled at the same time in state $\vec{e}$ and in state $\vec{e}_k$

$p_{ij} := \lambda[i, j](\vec{e}_k)$

If $(p_{ij} > 0)$ then

$\vec{e}_k := \vec{e}_k$

$e_k^i[i]--$

$e_k^i[j]++$

$\Delta_{ikj} := \Gamma[i, j](\vec{e}_k)$

Repeat for all pairs $(\vec{m}_l, p_l)$ in $\Delta_{ikj}$

$p_l := p_l \times p_{ij} \times p_k$

End Repeat

$\Delta_{ik} := \Delta_{ik} \cup \Delta_{ikj}$

End If

End Repeat

$p_k := p_k \times (1 - \sum_{j} p_{ij})$

$\Delta_i := \Delta_i \cup \Delta_{ik}$

End Repeat

$\Delta := \Delta \cup \Delta_i$

End Repeat

Return $\Delta$

End

Figure 3: Generation algorithm pseudocode
lost.

- **Output buffers**: there are \( N \) output buffers of size \( M \), each buffer being associated to an output link. The service time of each of the output buffers is a geometrical random variable of parameter \( r[i] \) (at each time slot there is a probability \( 1 - r[i] \) that the cell being served leaves the system).

- **Switching**: there is a single bus for packet switching. The bus works at a speed \( N \) times faster than the input links. We will take a bus service policy of type TDM (Time Division Multiplexing), in which each bus slot \( i \) is preassigned to input buffer \( i \); in this way, it is possible to process a cell from each of the input buffers in a single (arrival) time slot. Each of the cells arriving in input link \( i \) has as destination the output line \( j \) with probability \( D[i,j] \). If the cell being serviced at input buffer \( i \) has destination \( j \) and output buffer \( j \) is full, the cell will be blocked and will wait for its turn in the next cycle.

### 4.2 Performance evaluation

As input buffers only have a finite storage capacity, the cells arriving while they are full are lost. The loss probability \( P_L \) is one of the most important performance parameters in a packet switch. For each input link, we can compute the loss probability from the steady-state distribution by the following formula:

\[
P_L[i] = \frac{\sum_{\bar{n} \in F_P} \text{Prob}(\bar{n})(1 - q[i]) + \sum_{\bar{n} \in F_A} \text{Prob}(\bar{n})p[i]\alpha[i]}{\sum_{\bar{n} \in P} \text{Prob}(\bar{n})(1 - q[i]) + \sum_{\bar{n} \in A} \text{Prob}(\bar{n})p[i]\alpha[i]}
\]

with \( \text{Prob}(\bar{n}) \) the steady-state probability of the system being in state \( \bar{n} \), \( P \) the set of states where the input link is in passive state, \( A \) the set of states where the input link is in active state, \( F_P \) the set of states where the input link is in passive state and the input buffer is full, \( F_A \) the set of states where the input link is in active state and the input buffer is full.

The denominator is equal to the mean cell generation rate \( \rho[i] \) of input link \( i \), which verifies:

\[
\rho[i] = \frac{\alpha[i](1 - q[i])}{2 - p[i] - q[i]}.
\]
<table>
<thead>
<tr>
<th>$p$</th>
<th>0.10</th>
<th>0.30</th>
<th>0.50</th>
<th>0.70</th>
<th>0.90</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q$</td>
<td>0.82</td>
<td>0.86</td>
<td>0.90</td>
<td>0.94</td>
<td>0.98</td>
</tr>
<tr>
<td>$P_L$</td>
<td>$3.63 \times 10^{-4}$</td>
<td>$9.00 \times 10^{-3}$</td>
<td>$1.81 \times 10^{-3}$</td>
<td>$3.01 \times 10^{-3}$</td>
<td>$5.74 \times 10^{-3}$</td>
</tr>
<tr>
<td>Var(estimation)</td>
<td>$2.54 \times 10^{-9}$</td>
<td>$8.76 \times 10^{-9}$</td>
<td>$2.72 \times 10^{-8}$</td>
<td>$8.01 \times 10^{-8}$</td>
<td>$3.61 \times 10^{-7}$</td>
</tr>
</tbody>
</table>

Table 1: Simulation results for the cell loss probability in an ATM switch

The numerator can be computed from the steady-state distribution if the size of the system is small, or by regenerative simulation.

We will study the effect of the traffic characteristics on the loss probability $P_L = \sum_{i=1}^{N} P_L[i]$, for a given cell generation rate. We will take the following parameters as fixed: $N = 4$, $\alpha[i] = 0.5 \forall i$, $r[i] = 0.1 \forall i$, $D[i,j] = 1/4 \forall i,j$, and we will consider the values of $p[i]$ from 0.1 to 0.9 changing at the same time the values of $q[i]$ from 0.82 to 0.98 so that the generation rate is constant ($\rho[i] = 1/12$). The small values of $p[i]$ correspond to network traffic such that the average cell burst length is small; big values of $p[i]$ model the case where the average cell burst is long.

Using the BB library, we developed a balls and buckets model for the switch. The simulation facilities included in the library were employed to compute the cell loss probability. The computing equipment was a Sun Ultra 30 Model 295 workstation, and the total time needed to compute 100,000 independent runs for each of the five parameter combinations simulated was less than two hours. Table 1 shows the estimated values for cell loss probability (obtained by simulation), as well as the estimation precision. We can appreciate how the loss probability goes up very rapidly as the average burst length grows, even if the cell arrival rate doesn’t change. These phenomena are very important in ATM networks, and must be taken into account during the design phase, especially as traffic tends to be very bursty in the present-day telecommunications networks.

5 Conclusions

In this paper, we have presented an extension to the balls and buckets formalism allowing discrete time system modeling. This extension has been implemented in the BB library, giving access to the services it provides (Markov chain generation, storage, simulation).

As an application example, we have considered an ATM switch architecture and we have developed its balls and buckets model. This architecture performance was evaluated (for a number of different parameter values) by using the simulation primitives offered by the BB library; showing how the system performance, given a fixed load, depends on the traffic characteristics (if there are long bursts in the arrivals, the process autocorrelation is bigger, and performances diminish).

As future work, we can find on one hand the application of the BB library to other computer and communication systems evaluation. Another goal is to widen the scope of the library, by incorporating new model classes.

6 Acknowledgment

I would like to thank the anonymous referees from the PANEL conferences, for their helpful remarks.

I am also very grateful to Gerardo Rubino and all the MODEL project team at the Irisa, Rennes, France, where the first draft of this paper was prepared; without their kind reception and the many discussion opportunities they generated, this work would not have been possible.
References


