SRC TR 86-39

A Method for Handling Complex Markov Models of Distributed Algorithms

by

B.D. Plateau

TR-86-39

A METHOD FOR HANDLING COMPLEX MARKOV MODELS OF DISTRIBUTED ALGORITHMS

Brigitte D. Plateau

Computer Science Department

and

System Research Center

University of Maryland

College Park, Maryland 20842

This work was supported partially through NSF Grant NSFD CDR-85-00108.

Abstract: This paper is motivated by the study of the performance of distributed algorithms. The method presented here uses a representation of the algorithm as a network of state-transition graphs. The dynamic behavior of the algorithm is analyzed under Markovian assumptions. The state space explosion is handled by a decomposition technique. The generator of the chain is derived using tensor algebra operators.

1. Introduction

This work is motivated by the study of the performance of distributed algorithms. In the literature, there are many propositions for solving in a distributed way problems like mutual exclusion, termination detection, blocking detection or prevention, and information diffusion [Lamp78], [MeMu79], [RiAg81], [CaRo83], [DFVG83], [Rayn85], [ChLa85]. These algorithms should be examined for their correctness and compared for their performance. Measures like the number of message exchanges or whether or not the timestamp drift is bounded, can sometimes be derived from a rather simple analysis of the algorithm. In contrast, measures like response time, resource utilization and blocking probability have to take into account various complex timing patterns. These algorithms are characterized by two or more concurrent computations and synchronization constraints of various forms. To compare these algorithms, it is absolutely necessary to use the same type of assumptions for each distributed algorithm and a unified approach is essential to handle this comparison. Assuming a formal description of the algorithm is available in an algorithmic language, a simple way to keep a unified approach is to:

- exclude any behavior reduction,
- keep track of any relevant processes synchronization variables,
- have a uniform approach in terms of the random time distributions in the system.

This set of constraints forces the model to be very detailed, with a complex state space, and so is time consuming to solve. Once we have decided to work on a level of description very close to the application, one approach is to find the appropriate formalism and algorithms to handle this description. Indeed, any technique whose purpose is to describe parallelism - whether it is to achieve correctness proofs or to compute performance measures - has to be efficient regarding the rapid growth of the state space. An alternative to comparing these algorithms is to exhibit a set of rules that allows a systematic reduction of the model (in terms of the number of states) and leaves the performance measures invariant.

In this paper, a solution belonging to the first kind of approach is proposed. We present a graphical notation for describing the model that reflects naturally the processes composing the distributed algorithm. This representation leads to the identification of the synchronization constraints and to the decomposition of the transition rates of the related Markov model. Then Kronecker algebra is used to express the generator matrix of the Markov chain. This formulation is used for a more efficient computation of the performance measures.

The model is based on a state-transition graph description with timing information. A process in the distributed algorithm is represented by one or more state-transition graphs, which are related to each other. Typically, a node (also called a state) of one graph represents a specific state of process. An event is the transition from one state to another of one or more graphs. An event is represented by a zero-time transition: between events, the process spends time within states. This model is interpreted on a continuous time scale (a forthcoming paper will develop the same method on a discrete time scale). As a single graph represents the state of one process, additional information is used to express synchronization constraints among the processes.

The next example is a simple instance that shows situations that arise when modelling distributed algorithms [Plat85]. In Figure 1, we see three incomplete graphs that are intended to represent the behavior of a distributed algorithm. Graphs A_1 and A_2 have four states and are identical, while A_3 has only two states. Figure 1 represents only the states, without any transition rules. These transition rules will be added, step by step, through Figures 2 to 4. The global state space is the cross-product : $[1,4]\times[1,4]\times[1,2]$.(**)

In Figure 2, directed edges with strictly positive rates a_1, a_2 and a_3 are added, meaning that these particular transitions may fire independently with the corresponding rate. Notice that if all the transitions are of this type, each graph would be stochastically independent of the others, and thus could be studied separately.

In Figure 3, new directed edges appear with labels of the type: $c\ 1(E)$, where c is a transition rate and $1(E)$ is the characteristic function of a predicate E that has its arguments in the global state space .The function 1(E) has the value 1 if the predicate is true and 0 otherwise. The current state at instant t is an argument of the predicate. The transition takes place between times t and $t+dt$ with rate $c\ 1(E)$. This kind of dependency will be referred to as *a probabilistic dependency* among graphs and can be used in various situations:

(**) [1,4] represents the integer interval $\{1,2,3,4\}$.

- edges labeled with $b_1\ 1(A_3{=}1)$ and $b_1\ 1(A_3{\neq}1)$ are such that transition graph A_1 has a different dynamic behavior according to the current state of A_3.
- the edge labeled with $b_2\ 1(A_3{=}1)$ models a blocking situation: A_2 can only make the transition $2{\rightarrow}3$ if A_3 is in state 1. In this situation, it is important to have a careful interpretation of the model. As time proceeds, a predicate value may change. If it changes from 0 to 1, it means that the blocked activity is now allowed to start. If it changes from 1 to 0, it means that the current activity is stopped.

More generally, we may have transition rates that are any function of a set of parameters and the current global state.

In Figure 4, new labels of the type $c\,.(s)$, or (s), or $(s)\ p$ appear, where c is still a transition rate. A predicate whose argument is the current global state space may also appear in $c\ 1(E)\,(s)$. The symbol (s) stands for simultaneous, and all the transitions that have the symbol (s_i) with the same subscript i are involved in a simultaneous transition. For example, in transition (s_1) graphs A_1 and A_2 are involved, the starting state must be such that $(A_1{=}3,\ A_2{=}3)$, and they will proceed to transition $3{\rightarrow}4$ at the same instant, with rate c . This type of transition is called *a concurrent-firing* transition. Synchronization (s_2) is slightly more complex: it can take place only from a starting state where $(A_1{=}4,\ A_2{=}4,\ A_3{=}2)$. The transition of A_3 from 2 to 1 fires the simultaneous transitions of A_1 and A_2. However, the destination of A_2 depends on a routing probability p . We adopt the convention that firing rates precede the symbol (s), and routing probabilities follow this sign. Note that for graph A_3 there is a so-called *superposition* of different types of transitions out of state 2. This is denoted by the compound rate $d\ (s_2) + b_3$: The interaction involves a simultaneous transition with rate d for A_1, A_2 and A_3. If predicate $(A_1{=}4$ and $A_2{=}4)$ is not true, only the transition with rate b_3 can take place, involving only A_3. If this predicate is true, both transitions may take place, competing for the first place, and the first transition will disable the other. This type of transition is said to bear a *mixed* dependency.

In this example, we have seen the types of of dependencies that may occur:

(a) The transition is condition-free and fires at a fixed rate, or the transition has a rate and destination dependent on the global state space.

(b) The transition is bounded by a simultaneity constraint: the starting state of all the graphs involved in the dependency is defined by the set of arcs where the same label (s) appear. A unique transition rate is dedicated to the transition. This rate and each individual destination can also be dependent on the global state.

(c) Superpositions of the preceding dependencies may arise as it was seen in the example for (s_2).

These types of behavior can be identified easily in the framework of queuing networks: For example, two queues running in parallel, with independent inputs but service rates depending on the global state, is a case of probabilistic dependency between those two queues (Figure 5). In this example, compound rates M and N take into account probabilistic dependencies. On the other hand, if we consider two queues in tandem (Figure 6), concurrent-firing transitions occur at each termination of a service, which corresponds to an arrival for the other server. We have simultaneous jumps (s_1),...,(s_N), which are triggered by an end of service from server 1 or 2.

Modelling examples of distributed algorithms using this method can be found in [Plat84] and [PlTr86]. The goal of this paper is not to study and compare a class of algorithms, but to state carefully the theoretical basis of this approach. The behavior of the stochastic graph network is associated with a stochastic process (the designation *network* refers to the fact that synchronization constraints do exist between elementary graphs). This process clearly has the Markovian property since we assume that:

- any period of time has an exponential distribution,
- the constraints refer only to the current state and not the history of the process.

The Markov chain has several components, each representing the dynamic behavior of one state-transition graph. Due to synchronization constraints among components, it may be that none have a Markovian behavior in isolation. These components might have a finite or infinite discrete state space. The steady state behavior of the chain allows us to derive performance measures of interest.

Section 2 gives a formal definition for the different types of dependency between components of a vectorial Markov chain on continuous time and derive a basic decomposition theorem. Then section 3 gives the necessary background on Kronecker algebra to proceed. Finally, Section 4 shows how these formal dependencies can be translated in terms of building rules to derive the matrix of the Markov chain. The appendix applies the method to the examples of this introduction.

2. Dependencies between the components of a Markov chain

The idea in the example of Figures 1-4 was to build the model step by step, according to the different types of interactions that might occur among the model components. The result of this procedure is a graphic representation of a Markov chain as a network of stochastic graphs. To develop a mathematical definition of the interactions, we will work with the vectorial Markov chain representing the dynamic behavior of the network, rather than on the graph model. Considering the size of the problem, the first difficulty is to derive the transition matrix of this chain: As in any model of that type, the state space grows exponentially in terms of the degree of parallelism that is represented here by the number of components, which is also the number of elementary graphs. To solve this problem, the method presented here proposes to decompose the state space of the Markov chain so that each subspace can be identified by the single dependency it contains. This section defines the context in which such decompositions are possible.

Let us introduce some notations. $Z = (Z_t)_{t \in \mathbb{R}^+}$ is the homogeneous Markov chain with c components, $c > 1$, denoted by

$$Z_t = (X_{1,t}, \ldots, X_{c,t})$$

The set of *reachable* states of component X_i, $i \in [1,c]$, is E_i, which is a finite or numerable space. Its cardinality is denoted p_i, $p_i \in [1,+\infty]$. The chain Z is studied on the cross-product state space $E = E_1 \times \cdots \times E_c$. The set R denotes the set of reachable states for Z. In some cases R can be considerably smaller than E. But still, the idea is to keep the cross-product structure of E (which is the structure of independence) and attempt to make use of it instead of working on R, which has lost the regular product form structure of E. The set R is the irreducibility class we are interested in for performance evaluation. We assume that R is known.

The state space E is decomposed using canonical projections. For a subset $I = (i_1, \ldots, i_j)$ of $[1,c]$, $proj_I$ is the canonical projection of E on $E_I = E_{i_1} \times \cdots \times E_{i_j}$. Each set E_i is ordered and E possesses the corresponding lexicographical ordering. The set $proj_I(R)$ is denoted R_I. $\bar{I}$ is the complement of I in $[1,c]$ and $|I|$ is the number of elements in I. If $x = (x_1, \ldots, x_c) \in E$, we denote by $proj_I(x) = x_I = (x_{i_1}, \ldots, x_{i_j})$, and for the same subset I, $Z_I = (X_{i_1}, \ldots, X_{i_j})$. E_I is the cross-product space $E_{i_1} \times \cdots \times E_{i_j}$. Given two disjoint subsets I and K of $[1,c]$, and given x_I and y_K, we denote $\{x_I, y_K\}$ the vector $z = (z_j)_{j \in I \cup K}$ with $z_j = x_j$ if $i \in I$ and $z_i = y_i$ if $i \in K$. We emphasize that in any subvector, the ordering of components should be compatible with the natural ordering of $[1,c]$.

If I is a subset of $[1,c]$ and T is a subspace of R_I, we define

$$F(T) = \left\{ (x,x') \in R \times E \ \mid \ proj_I(x) \in T \ \text{ and } \ x \neq x' \right\}$$

$F(T)$ is the larger transition subspace such that for components with subscripts in I, the starting states of the transitions are restricted to T and the diagonal elements (of the form (x,x)) have been withdrawn.

The matrix that contains all transition rates is called the generator and is denoted Q_Z. We assume in the following that all the referenced limits exist and are denoted by:

$$(\forall (t,t') \in \mathbb{R}^+) \ \text{ with } \ t < t' \quad \text{and} \quad (\forall (y,y') \in R \times E)$$
$$\lim_{t' \to t} \frac{P(Z_{t'} = y' \mid Z_t = y)}{t' - t} = Q_Z(y,y') \quad \text{if} \quad y \neq y'$$
$$\lim_{t' \to t} \frac{1 - P(Z_{t'} = y' \mid Z_t = y)}{t' - t} = Q_Z(y,y') \quad \text{if} \quad y = y'$$

P is the corresponding probability measure and the initial probability distribution is such that: $(\forall y' \notin R)$ $(\forall y \in R)$ $P(Z_0 = y') = 0$ and $Q_Z(y,y') = 0$. $Q_Z(y',y)$ does not exist as the conditioning state is not reachable, but we arbitrarily set those elements to 0. Q_Z is a square matrix of dimension $|E|$.

In the next sections, we will define local properties of the chain based on the transition rates.

2.1. The probabilistic dependency

Probabilistic dependencies correspond to the intuitive idea presented in case (a) of the example in the introduction. In the sections that follow, x is used for a state that is fixed within the dependency, as y is a varying state. We want to emphasize that under the hypothesis of stochastic independence of all exponential variables, two periods end simultaneously with probability zero.

Definition 2.1: *Let $I = \{i\}$ be a single element subset of $[1,c]$, K another subset of $[1,c]$ that does not contain i and T_i a nonempty subspace of E_i. The component X_i is under the probabilistic dependency of Z_K on T_i if and only if*

$$(\forall (t,t') \in \mathbb{R}^+) \ \textit{ with } \ t < t' \quad \textit{and} \quad (\forall (y,y') \in F(T_i))$$

$$\lim_{t' \to t} \frac{P(\, Z_t{}' = y' \mid Z_t = y \,)}{t' - t} =$$

$$\lim_{t' \to t} \frac{P(\, Z_{i,t'} = y_i{}' \mid Z_{IUK,t} = y_{IUK} \,)}{t' - t} \times 1(\, y_{\bar{I}} = y_{\bar{I}}{}' \,) + \lim_{t' \to t} \frac{P(\, Z_{\bar{I},t'} = y_{\bar{I}}{}' \mid Z_t = y \,)}{t' - t} \times 1(\, y_i = y_i{}' \,)$$

In this definition, the dependency relationship between X_i and the rest of the chain has been identified: only components Z_K have an influence on the behavior of X_i and this dependency is expressed via the transition rate. This property holds whenever the view of the chain restricted to the subspace of transitions in $F(T_i)$. The assumption of stochastic independence of all the exponential delays, makes impossible any simultaneous jump of X_i and other components, when the starting state of X_i is T_i, which explains the exclusive predicates $1(\, y_{\bar{I}} = y_{\bar{I}}{}' \,)$ and $1(\, y_i = y_i{}' \,)$.

The restriction $(y, y') \in F(T_i)$ excludes diagonal elements, which require a different formula. To avoid heavier notations, those are omitted. They can be systematically deduced from the fact that the sum of each row of Q_Z is zero.

If the subset K is empty, component X_i has a local independence on subspace T_i. Under the assumptions of this definition, for any subspace $T_i{}' \subset T_i$, X_i is also under the probabilistic influence of Z_K on $T_i{}'$. As our goal is to decompose the transition space, and to achieve this with a minimum number of operations, we try to produce subspaces with a property of *maximality.* As we can notice in many examples, such a decomposition is not unique.

A probabilistic influence γ will be represented by the triple $(I_\gamma = \{i_\gamma\}, K_\gamma, T_{i_\gamma})$ and the set of coefficients:

$$(\forall (y, y') \in F(T_{i_\gamma}))$$

$$c_\gamma(y, y') = \lim_{t' \to t} \frac{P(\, X_{i_\gamma t'} = y_{i_\gamma}{}' \mid Z_{I_\gamma UK_\gamma t} = y_{I_\gamma UK_\gamma} \,)}{t' - t} \times 1(y_{\bar{I}_\gamma} = y_{\bar{I}_\gamma}{}' \,) \qquad 2.1$$

so that the global transition rate can be expressed:

$$(\forall (y, y') \in F(T_{i_\gamma})) \qquad \lim_{t' \to t} \frac{P(\, Z_t{}' = y' \mid Z_t = y \,)}{t' - t} \qquad 2.1.1$$

$$= c_\gamma(y, y') + \lim_{t' \to t} \frac{P(\, Z_{\bar{I}_\gamma t'} = y_{\bar{I}_\gamma}{}' \mid Z_t = y \,)}{t' - t} \times 1(y_{i_\gamma} = y_{i_\gamma}{}' \,)$$

Component i_γ is called the *moving component* of this dependency. This summation involving exclusive predicate functions is the type of decomposition that is searched for in the next two sections. The result should lead to a decomposition of the transition rates of Q_Z into sums where each factor is clearly identified.

2.2. The concurrent-firing dependency

We now define the theoretical background for the type of dependency referred to as type (b) in the introduction, where two groups of components can be defined. The first group is composed of those components involved in the simultaneous jump. These components are still called *moving components.* Furthermore, the transition probabilities and each individual destination are dependent on the current state of another group of components called the *interacting components.* These transition probabilities and destinations are also naturally dependent on the state of the moving components as we are under Markovian assumptions. Thus, by definition, interacting components are non-moving components. This can be formalized as follows.

Definition 2.2 : *Let I and K be two disjoint subsets of $[1,c]$, $|I|>1$, $T_I=\{x_I\}$ a singleton of R_I and p a mapping from $[(T_I\times E_K)\cap R_{I\cup K}]\times E_I$ to $\mathbb{R}^+$. Then we say the components Z_I are bounded by a concurrent-firing dependency on T_I under the probabilistic influence of Z_K and driven by the mapping p if and only if*

(i) $(\forall (y,y') \in F(T_I))$

$$\lim_{t'\to t} \frac{P(Z_{t'}=y' \mid Z_t=y)}{t'-t} =$$

$$p(x_I, y_K, y_I') \times 1(y_{\bar I}=y_{\bar I}') + \lim_{t'\to t} \frac{P(Z_{\bar I,t'}=y_{\bar I}' \mid Z_t=y)}{t'-t} \times 1(y_I'=y_I)$$

(ii) $(\forall J\subset I)$ $J\neq I$ $J\neq\emptyset$ $(\forall (y,y') \in F(proj_J(T_I)))$ *such that* $(\forall j\in I-J)$ $y_j\neq x_j$

$$\lim_{t'\to t} \frac{P(Z_{t'}=y' \mid Z_t=y)}{t'-t} = \lim_{t'\to t} \frac{P(Z_{\bar J,t'}=y_{\bar J}' \mid Z_t=y)}{t'-t} \times 1(y_J'=y_J)$$

provided that mapping p has the following property: If we denote $f_I(y_K)$ the set $\{x_I' \in E_I \mid x_I'\neq x_I, \text{and } p(x_I,y_K,x_I')\neq 0\}$, then we require that all elements of $f_I(y_K)$ are such that $(\forall i\in I)$ $proj_i(x_I')\neq proj_i(x_I)$.

Z_I are the moving components and Z_K the interacting ones. This definition states that the simultaneous transition is the only way for Z_I to jump out of $T_I=\{x_I\}$. No simultaneous jump may occur involving both X_I and $X_{\bar I}$ on T_I. Case (i) represents the fact that the condition on the starting state of this concurrent-firing dependency is fulfilled with $y_I = x_I$, as $(y, y') \in F(T_I)$. This notion of enforced simultaneity is represented by the single transition rate $p(x_I, y_K, y_I')$. This transition rate is different from 0 only when all components are moving, which is expressed by the hypothesis on all elements in $f_I(y_K)$: $(\forall i \in I)\quad proj_i(x_I') \neq proj_i(x_I)$. $f_I(y_K)$ is the set of possible destinations of this simultaneous jump under the condition $Z_K = y_K$. In case (i), either the simultaneous jump can take place, or none of the components in I can move. In case (*ii*), only part of the components (namely in J) are ready for the simultaneous jump and those cannot move at this instant. A parametric influence exists in the sense that the transition rate mapping p only depends on x_I, y_K and y_I', meaning that we have identified the interacting components Z_K. The case where I has a single element is excluded as it would turn a concurrent-firing dependency into a simple probabilistic one and we want these designations to be exclusive. There might be concurrent-firing dependencies where the subset K is empty and the parametric influence disappears leaving only the synchronization constraint. A particular case occurs when $\overline{I} = \emptyset$. In this case we assume $\lim_{t' \to t} \dfrac{P(Z_{\overline{I},t'} = y_{\overline{I}}' \mid Z_t = y)}{t' - t} = 0$. This holds for the rest of the paper.

In this definition, T_I is restricted to a singleton as complex starting state set can be decomposed as a collection of those elementary concurrent-firing dependencies. Each concurrent-firing dependency λ is defined by the 5-tuple $(I_\lambda, K_\lambda, T_{I_\lambda}, p_\lambda, f_{I_\lambda})$ and a set of coefficients $c_\lambda(y, y')$ such that

$$(\forall (y, y') \in F(T_{I_\lambda}))$$
$$c_\lambda(y, y') = \sum_{x_{I_\lambda}' \in f_{I_\lambda}(y_{K_\lambda})} p(x_{I_\lambda}, y_{K_\lambda}, y_{I_\lambda}') \times 1(y_{I_\lambda}' = x_{I_\lambda}') \times 1(y_{\overline{I}_\lambda} = y_{\overline{I}_\lambda}') \qquad 2.2(i)$$

and $(\forall J \subset I_\lambda)\ J \neq I_\lambda\quad J \neq \emptyset\quad (\forall (y, y') \in F(proj_J(T_I)))$ such that $(\forall j \in I_\lambda - J)\ y_j \neq x_j$
$$c_\lambda(y, y') = 0 \qquad 2.2(ii)$$

Under the conditions of the theorem and with H_λ being I_λ or J_λ, respectively, we have

$$\lim_{t' \to t} \frac{P(Z_t' = y' \mid Z_t = y)}{t' - t} \qquad 2.2.1$$
$$= c_\lambda(y, y') + \lim_{t' \to t} \frac{P(Z_{\overline{H}_\lambda, t'} = y_{\overline{H}_\lambda}' \mid Z_t = y)}{t' - t} \times 1(y_{H_\lambda} = y_{H_\lambda}')$$

which is again the summation form we are looking for.

2.3. The mixed dependency

If we summarize the preceding discussion, a probabilistic dependency is related to a dependency via transition rates for a single component and the concurrent-firing dependency is related to a synchronization constraint. Nevertheless, there might be cases where a stochastic dependency and a concurrent-firing dependency are superimposed as we have seen in the introductory example. Multiple superpositions of this type can be reduced to a single one by increasing the sets I and K.

Definition 2.3: *Let I and K be two disjoint subsets of $[1,c]$ with $|I|>1$, I_0 a nonempty subset of I, $T_I=\{x_I\}$ a singleton of R_I and p a mapping from $[(T_I \times E_K)\cap R_{IUK}]\times E_I$ to $\mathbb{R}^+$. Then we say the components Z_I are bounded by a mixed dependency on T_I under the probabilistic influence of Z_K and driven by the mapping p if and only if*

$$(i)\ (\forall (y,y') \in F(T_I))$$

$$\lim_{t'\to t} \frac{P(Z_{t'}=y' \mid Z_t=y)}{t'-t} = p(x_I, y_K, y_I') \times 1(y_{\bar{I}}=y_{\bar{I}}')$$

$$+ \lim_{t'\to t} \frac{P(X_{i,t'}=y_i' \mid Z_{IUK,t}=y_{IUK})}{t'-t} \times 1(i \in I_0) \times 1(y_i \neq y_i') \times \prod_{k\neq i} 1(y_k=y_k')$$

$$+ \lim_{t\to t'} \frac{P(Z_{\bar{I},t'}=y_{\bar{I}}' \mid Z_t=y)}{t'-t} \times 1(y_I'=y_I)$$

$$(ii)\ (\forall J\subset I)\ J\neq I\ \ J\neq\emptyset\ \ (\forall (y,y') \in F(proj_J(T_I)))\ \text{such that}\ (\forall j \in I-J)\ y_j\neq x_j$$

$$\lim_{t'\to t} \frac{P(Z_{t'}=y' \mid Z_t=y)}{t'-t} =$$

$$\lim_{t'\to t} \frac{P(Z_{\bar{J},t'}=y_{\bar{J}}' \mid Z_t=y)}{t'-t} \times 1(y_J'=y_J)$$

$$+ \lim_{t'\to t} \frac{P(X_{i,n}=y_i' \mid Z_{IUK,n-1}=y_{IUK})}{t'-t} \times 1(i \in I_0\cap J) \times 1(y_i\neq y_i') \times \prod_{k\neq i} 1(y_k=y_k')$$

provided that mapping p has the following property: If we denote $f_I(y_K)$ the set $\{x_I' \in E_I \mid x_I' \neq x_I, \text{and } p(x_I,y_K,x_I')\neq 0\}$, then we require that all elements of $f_I(y_K)$ are such that $(\forall i \in I)$ $proj_i(x_I')\neq proj_i(x_I)$.

This definition essentially says that, from the starting state x_I exactly one of the following may occur: a simultaneous jump, a single jump of one component in I_0, or any jump that does not involve components in I. The superposition in the mixed dependency appears in the summation of two types of coefficients: a single coefficient for the synchronized event, which involve all components in I, or local rates for the probabilistic dependencies, which involves a single component in I_0. Case (i) reflects the fact that the starting condition $x_I = y_I$ is satisfied, from which transitions may occur for the simultaneous jump or any single jump of components in I_0. In case (ii), the starting condition is not fulfilled and the components in $(I - I_0) \cap J$ are blocked, as those in $I_0 \cap J$ can move according to their probabilistic dependency. The diagonal elements are still omitted.

Each mixed dependency ϕ is defined by the 6-tuple $(I_\phi, K_\phi, I_{0_\phi}, T_{I_\phi}, p_\phi, f_{I_\phi})$ and a set of coefficients $c_\phi(y, y')$ such that

$$(\forall (y, y') \in F(T_{I_\phi}))$$

$$c_\phi(y, y') = \sum_{x_{I_\phi}' \in f_{I_\phi}(y_{K_\phi})} p(x_{I_\phi}, y_{K_\phi}, y_{I_\phi}') \times 1(y_{I_\phi}' = x_{I_\phi}') \times 1(y_{\bar{I}_\phi} = y_{\bar{I}_\phi}') \qquad 2.3(\mathrm{i})$$

$$+ \sum_{i \in I_0} \lim_{t' \to t} \frac{P(X_{i,t'} = y_i \mid Z_{IUK,t} = y_{IUK})}{t' - t} 1(y_i \neq y_i') \times \prod_{k \neq j} 1(y_k = y_k')$$

and $(\forall J \subset I_\phi)\ J \neq I_\phi\ \ J \neq \emptyset\ \ (\forall (y, y') \in F(proj_J(T_I)))$ such that $(\forall j \in I_\phi - J)\ \ y_j \neq x_j$

$$c_\phi(y, y') = \qquad 2.3(\mathrm{ii})$$

$$\sum_{i \in I_0 \cap J} \left[\lim_{t' \to t} \frac{P(X_{i,t'} = y_i' \mid Z_{IUK,t} = y_{IUK})}{t' - t} 1(y_i \neq y_i') \times \prod_{k \neq j} 1(y_k = y_k') \right]$$

So, we can write the following summation formula, with H_ϕ being I_ϕ or J_ϕ, respectively

$$(\forall (y, y') \in F(T_{H_\phi})) \qquad \lim_{t' \to t} \frac{P(Z_t' = y' \mid Z_t = y)}{t' - t} \qquad 2.3.1$$

$$= c_\phi(y, y') + \lim_{t' \to t} \frac{P(Z_{\bar{H}_\phi t'} = y_{\bar{H}_\phi}' \mid Z_t = y)}{t' - t} \times 1(y_{H_\phi} = y_{H_\phi}')$$

As is clear from these definitions, probabilistic, concurrent-firing and mixed dependencies show an increasing complexity. As a matter of fact, a concurrent-firing dependency could have been a particular case of mixed dependency: but the designation mixed is used when it is neither a concurrent-firing dependency nor a probabilistic dependency, as we assume $|I| > 1$ and $I_0 \neq \emptyset$.

In the area of distributed algorithm modelling , this method will be tractable if the model has only a few concurrent-firing and a very few mixed dependencies or if they show some regularities such as linearity (as in the queuing examples in the introduction). Moreover, we will restrict ourselves to these three definitions. It is clear that they do not cover all dependencies that may arise among components of a Markov chain. We have defined a sort of *pure* probabilistic dependency and a *pure* concurrent-firing dependency and started to consider one compound structure. It is possible to think of more complex patterns, such as an embedded scheme of synchronization constraints. It is not clear whether such situations deserve another definition or should be seen as a recursive application of the same basic rules. This will be the subject of future work. For the time being, we restrict ourselves to problems where these three basic rules lead to an exhaustive analysis of the stochastic behavior of each component X_i of the chain. The appendix shows how this method applies to each of the examples of the introduction.

2.4. Partitioning of the state space

The line of thought is to isolate the parts of the Markov chain having the same properties. Following the idea of the previous sections, isolation is made using a partition of the state space for each component or even group of components when they cannot be separated by a synchronization constraint. The partition of E_i is such that

- if X_i is under a probabilistic dependency from Z_K on T_i, then T_i is denoted $T_i^{m_i}$ as the m_i-th member of this partition,
- if X_i is a moving component within a concurrent-firing or mixed dependency π on T_{I_π}, then $T_i^{m_i} = proj_i(T_{I_\pi})$ is a member of this partition.

The sets $T_i^{m_i}$ cannot be overlapping as the Definitions 2.1, 2.2 and 2.3 are mutually exclusive, so one component can only be involved in one dependency on a subspace $T_i^{m_i}$. The set of $(T_i^{m_i})_{m_i \in [1,t_i]}$ covers E_i, as we assumed a problem that shows only these three types of dependencies. A partition of the global state space is obtained with all the cross-products subspaces

$$T_{\overline{m}} = \prod_{i=1}^{c} T_i^{m_i} \qquad \overline{m} = (m_1, \ldots, m_c) \quad \text{and} \quad 1 \le m_i \le t_i$$

$T_{\overline{m}}$ is called a *basic paving block* of this problem. This decomposition, which is not unique, leads to the decomposition of the global transition rates using these partial views corresponding to dependencies. The interest of the whole technique lies in the belief that it will enable us to study the behavior of Z by looking at small groups of components over subspaces of transitions where their stochastic properties can be clearly identified and practically handled.

2.5. The decomposition theorem

From an exhaustive analysis of the chain, we withdraw $\Pi\Gamma$, the set of all probabilistic dependencies, $\Pi\Lambda$, the set of all concurrent-firing dependencies and $\Pi\Phi$, the set of all mixed dependencies. The next definition helps to classify these dependencies according to the paving blocks.

Definition 2.5: *A basic paving block $T_{\overline{m}}$ and a dependency π defined on $T_{I_\pi}=\{x_{I_\pi}\}$ are consistent if and only if $proj_{I_\pi}(T_{\overline{m}}) = T_{I_\pi}$. If this dependency π and the paving block $T_{\overline{m}}$ are inconsistent, we define the biggest set $J \subset I_\pi$ that has the property $proj_J(T_{I_\pi})=proj_J(T_{\overline{m}})$, as being the trace of π on $T_{\overline{m}}$. If $J = \emptyset$, π is said to be totally inconsistent with $T_{\overline{m}}$. Otherwise π is said to be partially consistent with $T_{\overline{m}}$.*

The partial consistency is related to cases (ii) of Definitions 2.1 and 2.2, and consistency to case (i). We notice that the notion of consistency is only related to moving components. Considering any element y in R, there exists a unique paving block $T_{\overline{m}}$ that contains y. Related to $T_{\overline{m}}$, Γ, Λ, Λ', Φ and Φ' are respectively the list of consistent probabilistic dependencies, the lists of the consistent and partially consistent concurrent-firing dependencies and the lists of the consistent and partially consistent mixed dependencies. These dependencies define respectively the sets of components $(I_\gamma)_{\gamma\in\Gamma}$, $(I_\lambda)_{\lambda\in\Lambda}$, $(J_\lambda)_{\lambda\in\Lambda'}$, $(I_\phi)_{\phi\in\Phi}$, $(J_\phi)_{\phi\in\Phi'}$ (where J is the trace as defined in Definition 2.5), which altogether form a partition of $[1,c]$. This property is important for the following. The next theorem gives a decomposition formula valid for any paving block $T_{\overline{m}}$.

Theorem 2.1: *Given the dependencies $\Pi\Gamma$, $\Pi\Lambda$, and $\Pi\Phi$ from an exhaustive analysis of the chain Z, given a paving block $T_{\overline{m}}$ and y in $T_m \cap R$, and the set of dependencies Γ, Λ, Λ', Φ and Φ' defined above, we have:*

$$(\forall y' \in E) \qquad y \neq y' \qquad Q_Z(y,y') = \tag{2.5}$$

$$\sum_{\gamma\in\Gamma} c_\gamma(y,y') + \sum_{\lambda\in\Lambda} c_\lambda(y,y') + \sum_{\lambda'\in\Lambda'} c_{\lambda'}(y,y') + \sum_{\phi\in\Phi} c_\phi(y,y') + \sum_{\phi'\in\Phi'} c_{\phi'}(y,y')$$

where the coefficients c_γ, $c_\lambda, c_{\lambda'}$, c_ϕ *and* $c_{\phi'}$ *are those defined, for each dependency, in equations 2.1, 2.2 (i), 2.2 (ii), 2.3 (i) and 2.3 (ii) respectively.*

This proof is based on an iterative use of equations 2.1.1, 2.2.1 and 2.3.1. The use of these coefficients is consistent with the definitions as $(\forall i \in [1,c])$ if $(y,y') \in (T_{\overline{m}} \cap R) \times E$ and $y \neq y'$, then $(y,y') \in F(T_{I_\pi})$ for all consistent dependencies π and $(y,y') \in F(T_{J_\pi})$ for any partially consistent dependency whose trace on $T_{\overline{m}}$ is J_π. Considering the null values of some of these coefficients, equation 2.5 becomes

$$Q_Z(y,y') = \sum_{\gamma\in\Gamma} c_\gamma(y,y') + \sum_{\lambda\in\Lambda} c_\lambda(y,y') + \sum_{\phi\in\Phi} c_\phi(y,y') + \sum_{\phi'\in\Phi'} c_{\phi'}(y,y') \quad 2.6$$

The proof must derive a sum decomposition of a global transition rate, according to the local dependencies.

Proof: We assume for instance that Γ is nonempty and for any $\gamma \in \Gamma$ we have from equation 2.1.1

$$(\forall y \in T_{\overline{m}} \cap R)\ (\forall y' \in E)\ \ y \neq y' \qquad 2.7$$

$$Q_Z(y,y') = c_\gamma(y,y') + \lim_{t' \to t} \frac{P(Z_{\overline{I}_\gamma,t'} = y_{\overline{I}_\gamma}' \mid Z_t = y)}{t'-t} \times 1(y_{I_\gamma} = y_{I_\gamma}')$$

To start with an inductive argument, we assume we have a subset of dependencies Δ defined by

$$\Delta = \Gamma_0 \cup \Lambda_0 \cup \Phi_0 \cup \Phi_0' \quad \text{with } \Gamma_0 \subset \Gamma,\ \Lambda_0 \subset \Lambda,\ \Phi_0 \subset \Phi \text{ and } \Phi_0' \subset \Phi'$$

J is the set of moving components for all dependencies of Δ. We assume we have the following equality for all $y \in T_{\overline{m}} \cap R$, $y' \in E$ and $y \neq y'$

$$Q_Z(y,y') = \sum_{\delta\in\Delta} c_\delta(y,y') + \lim_{t' \to t} \frac{P(Z_{\overline{J},t'} = y_{\overline{J}}' \mid Z_t = y)}{t'-t} \times 1(y_J = y_J') \qquad 2.8$$

Then for any dependency ν which is not in Δ, we have

$$Q_Z(y,y') = c_\nu(y,y') + \lim_{t' \to t} \frac{P(Z_{\overline{I}_\nu,t'} = y_{\overline{I}_\nu}' \mid Z_t = y)}{t'-t} \times 1(y_{I_\nu} = y_{I_\nu}') \qquad 2.9$$

On the other hand we have a decomposition of the residual term in 2.8

$$\lim_{t' \to t} \frac{P(Z_{\overline{J},t'} = y_{\overline{J}}' \mid Z_t = y)}{t'-t} \qquad 2.10$$

$$= \sum_{y_J' \in E_J} \lim_{t' \to t} \frac{P(Z_{J,t'} = y_J' \ , Z_{\bar{J},t'} = y_{\bar{J}}' \ \mid Z_t = y)}{t' - t}$$

If we call H the set of moving components involved in the set of dependencies $\Delta' = \Delta U\{\nu\}$, then $H = JUI_\nu$ and we can apply 2.9 to each term of sum 2.10, yielding

$$\lim_{t' \to t} \frac{P(Z_{\bar{J},t'} = y_{\bar{J}}' \ \mid Z_t = y)}{t' - t} \tag{2.11}$$

$$= \sum_{y_J' \in E_J} \left[c_\nu(y, y') + \lim_{t' \to t} \frac{P(Z_{J,t'} = y_J' \ , Z_{\bar{H},t'} = y_{\bar{H}}' \ \mid Z_t = y)}{t' - t} \times 1(y_{I_\nu} = y_{I_\nu}') \right]$$

In this sum, there is only one term where $c_\nu(y, y')$ is not null (namely the term corresponding to the nonzero value of the function $1(y_{\bar{I}_\nu} = y_{\bar{I}_\nu}')$, knowing that $J \subset I_\nu$), and we have

$$\lim_{t' \to t} \frac{P(Z_{\bar{J},t'} = y_{\bar{J}}' \ \mid Z_t = y)}{t' - t} \tag{2.12}$$

$$= c_\nu(y, y') + \sum_{y_J' \in E_J} \lim_{t' \to t} \frac{P(Z_{J,t'} = y_J' \ , Z_{\bar{H},t'} = y_{\bar{H}}' \ \mid Z_t = y)}{t' - t} \times 1(y_{I_\nu} = y_{I_\nu}')$$

At this point, we can process the $\sum$, so that from 2.8 and 2.12, and knowing that $c_\nu(y, y') = 1(y_J = y_J') \times c_\nu(y, y')$ and $1(y_{I_\nu} = y_{I_\nu}') \times 1(y_J = y_J') = 1(y_H = y_H')$, we deduce

$$Q_Z(y, y') = \sum_{\delta \in \Delta'} c_\delta(y, y') + \lim_{t' \to t} \frac{P(Z_{\bar{H},t'} = y_{\bar{H}}' \ \mid Z_t = y)}{t' - t} \times 1(y_H = 1 y_H') \tag{2.13}$$

With this iterative procedure, by scanning all dependencies in Γ, Λ, Λ' , Φ and Φ' , we reach the point where the set $\bar{H}$ is reduced to an empty set, as the sets $(I_\gamma)_{\gamma \in \Gamma}$, $(I_\lambda)_{\lambda \in \Lambda}$, $(J_\lambda)_{\lambda \in \Lambda'}$, $(I_\phi)_{\phi \in \Phi}$, $(J_\phi)_{\phi \in \Phi'}$ form a partition of $[1,c]$. This concludes the proof.

This theorem closes the section dedicated to the stochastic analysis of the chain. We know that within the framework of this model, we can obtain a decomposition of the global transition rates into a sum of transition rates, which are of a more elementary type. Intuitively, cuts are made according to local stochastic independence. The following will focus on the other aspect, namely developing matrix tools that correspond to this stochastic analysis. For this, the next section gives a small introduction to the matrix tools necessary for further work.

3. Elements of Kronecker algebra and their extensions

This section states the definitions and gives some information on the Kronecker algebra utilization [Maza83]. Let $M(n)$ be the set of square matrices of size n with elements in $\mathbb{R}$. On the one hand, a matrix can be represented by $A = (a(i,j))_{1 \le i,j \le n}$. On the other hand, a matrix in $M(np)$ can be represented by the set of its terms, according to their positions in the n^2 blocks in $M(p)$:

$$C \in M(np) \quad C = (c(\overline{i},\overline{j}))_{\overline{i},\overline{j} \in L(n,p)}$$
$$\text{with} \quad L(n,p) = \{\overline{i} \mid \overline{i} = (i_1,i_2) \ \ 1 \le i_1 \le n, \ 1 \le i_2 \le p\}$$

The block coordinate of element $(\overline{i},\overline{j})$ is (i_1,j_1), and (i_2,j_2) is its situation within this block. With a direct generalization procedure, the elements of matrix $C \in M(p_1 p_2 \cdots p_c)$ are referenced as follows:

$$C = (c(\overline{i},\overline{j}))_{\overline{i},\overline{j} \in L(p_1,\ldots,p_c)}$$
$$\text{with} \quad L(p_1,\ldots,p_c) = \{\overline{i} \mid \overline{i} = (i_1,\ldots,i_c) \text{ and } (\forall k \in [1,c]) \quad 1 \le i_k \le p_k\}$$

$L(p_1, \ldots, p_c)$ has the lexicographical ordering. We denote $\overline{p} = (p_1,\ldots,p_c)$, and for $\overline{i} \in L(\overline{p})$ define the number

$$s(\overline{i}) = \prod_{j=1}^{c} i_j$$

$s(\overline{p})$ is the cardinality of $L(\overline{p})$. $\overline{i_k}$ is the vector obtained from $\overline{i}$ by removing the i_k component. $L_k(\overline{p})$ is the set of all $\overline{i_k}$ and $s(\overline{i_k})$ is the cardinality of $L_k(\overline{p})$. We define also

$$r_k(\overline{i}) = \prod_{j=1}^{k-1} i_j \qquad \text{and} \qquad t_k(\overline{i}) = \prod_{j=k+1}^{c} i_j$$

where any empty product is equal to 1. Id_p is the identity matrix in $M(p)$.

Definition 3.1: *The Kronecker product and sum of matrix A in M(n) and B in M(p) are matrices C and D in M(np) defined by:*

$$C = A \otimes B$$
$$\textit{with} \quad c(\overline{i},\overline{j}) = a(i_1,j_1)b(i_2,j_2) \quad \textit{and} \quad \overline{i} = (i_1,i_2)\ ,\ \overline{j} = (j_1,j_2)$$

$$\textit{and} \quad D = A \oplus B = A \otimes Id_p + Id_n \otimes B$$

$$\textit{with} \quad d(\overline{i},\overline{j}) = a(i_1,j_1) \quad \textit{if} \quad i_1 \ne j_1 \quad \textit{and} \quad i_2 = j_2$$
$$\textit{with} \quad d(\overline{i},\overline{j}) = b(i_2,j_2) \quad \textit{if} \quad i_1 = j_1 \quad \textit{and} \quad i_2 \ne j_2$$

$$\textit{with} \quad d(\bar{i},\bar{j}) = -a(i_1,j_1) - b(i_2,j_2) \quad \textit{if} \quad i_1 = j_1 \quad \textit{and} \quad i_2 = j_2$$
$$\textit{with} \quad d(\bar{i},\bar{j}) = 0 \quad \textit{otherwise}$$

As an example consider,

$$A = \begin{bmatrix} a(1,1) & a(1,2) \\ a(2,1) & a(2,2) \end{bmatrix} \qquad B = \begin{bmatrix} b(1,1) & b(1,2) \\ b(2,1) & b(2,2) \end{bmatrix}$$

Then

$$A \otimes B = \begin{bmatrix} a(1,1)b(1,1) & a(1,1)b(1,2) & a(1,2)b(1,1) & a(1,2)b(1,2) \\ a(1,1)b(2,1) & a(1,1)b(2,2) & a(1,2)b(2,1) & a(1,2)b(2,2) \\ a(2,1)b(1,1) & a(2,1)b(1,2) & a(2,2)b(1,1) & a(2,2)b(1,2) \\ a(2,1)b(2,1) & a(2,1)b(2,2) & a(2,2)b(2,1) & a(2,2)b(2,2) \end{bmatrix}$$

and

$$A \oplus B = \begin{bmatrix} a(1,1)+b(1,1) & b(1,2) & a(1,2) & 0 \\ b(2,1) & a(1,1)+b(2,2) & 0 & a(1,2) \\ a(2,1) & 0 & a(2,2)+b(1,1) & b(1,2) \\ 0 & a(2,1) & b(2,1) & a(2,2)+b(2,2) \end{bmatrix}$$

Notice the embedded block structure of the result $A \otimes B$ and $A \oplus B$. It is clear from this last definition that the operators $\otimes$ and $\oplus$ are not commutative.

It can be easily shown that if X is a Markov chain (indexed by a discrete time scale) with transition matrix A, and Y is another Markov chain independent from the first with transition matrix B, the vector (X,Y) has the transition matrix $A \otimes B$. Similarly, if X is a Markov chain indexed by a continuous time scale, with generator matrix A, and Y is another Markov chain independent from the first one with generator matrix B, the vector (X,Y) has the generator matrix $A \oplus B$. We see that stochastic independence has a straightforward representation in terms of Kronecker product and sum. Note that in the first case the Kronecker product is the natural operator, while in the second it is the Kronecker sum. Nevertheless, we restrict ourselves to the product operator. This is possible because the sum, which we exclude, can be expressed as products (as in the definition). Complementary information can be found in [Davi81] concerning Kronecker algebra properties.

The classical Kronecker product has simple matrices as arguments. We generalize it for vector arguments. Given $(A^i)_{1 \le i \le p}$, p matrices in $M(n)$, we denote $\bar{A}$ the vector of matrices $\bar{A} = (A^i)_{1 \le i \le p}$, which is considered as an element of $M(n)^p$.

Definition 3.2 : *Given the vectors of matrices* $\overline{A} = (A^i)_{1 \le i \le p}$, *where* $A^i \in M(n)$, *and* $\overline{B} = (B^i)_{1 \le i \le n}$, *where* $B^i \in M(p)$, *the generalized Kronecker product of* $\overline{A}$ *and* $\overline{B}$ *is the matrice* C *in* $M(np)$ *defined by:*

$$C = \overline{A} \otimes \overline{B}$$

$$\text{with} \quad c(\overline{i},\overline{j}) = a^{i_2}(i_1,j_1) \, . \, b^{i_1}(i_2,j_2) \quad \text{if } \overline{i} = (i_1,i_2) \quad , \quad \overline{j} = (j_1,j_2)$$

$$\text{and} \quad A^i = (a^i(k,l))_{1 \le k,l \le n} \quad , \quad B^i = (b^i(k,l))_{1 \le k,l \le p}$$

For example, if $n = p = 2$

$$\overline{A} \otimes \overline{B} = \begin{bmatrix} a^1(1,1)b^1(1,1) & a^1(1,1)b^1(1,2) & a^1(1,2)b^1(1,1) & a^1(1,2)b^1(1,2) \\ a^2(1,1)b^1(2,1) & a^2(1,1)b^1(2,2) & a^2(1,2)b^2(2,1) & a^2(1,2)b^1(2,2) \\ a^1(2,1)b^2(1,1) & a^1(2,1)b^2(1,2) & a^1(2,2)b^2(1,1) & a^1(2,2)b^2(1,2) \\ a^2(2,1)b^2(2,1) & a^2(2,1)b^2(2,2) & a^2(2,2)b^2(2,1) & a^2(2,2)b^2(2,2) \end{bmatrix}$$

The terms in the first row of this matrix come from A^1 and B^1, those of the second row from A^2 and B^1, those of the third from A^1 and B^2, these of the fourth from A^2 and B^2. If all the matrices A^i are equal to A and all the matrices B^i are equal to B, then $\overline{A} \otimes \overline{B} = A \otimes B$. By an extension of notation, if A is a matrix in $M(n)$, $\overline{A}$ is the vector $(A)_{1 \le i \le p}$ whose components are all equal to A, and $\overline{B} = (B^i)_{1 \le i \le n}$, then we write:

$$A \otimes \overline{B} = \overline{A} \otimes \overline{B}$$

For A in $M(n)$, we denote $l_i(A)$ the matrix of the same dimension whose i-th row is equal to the i-th row of A, the other terms being zero. Notice that for every matrix A in $M(n)$ and B in $M(p)$ the following equality holds:

$$l_i(A) \otimes l_j(B) = l_{\overline{i}}(A \otimes B) \quad \text{with} \quad \overline{i} = (i,j)$$

This property can be generalized as follows: If $\overline{A} \in M(n)^p$ and $\overline{B} \in M(p)^n$ then

$$\overline{A} \otimes \overline{B} = \sum_{i=1}^{n} \sum_{j=1}^{p} l_i(A^j) \otimes l_j(B^i)$$

This equality is very useful to reduce formulas where generalized Kronecker products occur into expressions with classical Kronecker products only. A particular case is:

$$\overline{A} \otimes B = \sum_{j=1}^{p} A^j \otimes l_j(B)$$

$$A \otimes \overline{B} = \sum_{i=1}^{n} l_i(A) \otimes B^i$$

These identities are often used to simplify the generalized Kronecker sums. They can also help in defining generalized Kronecker operations with more than two arguments, as we will see in the next definitions.

It is easy to check that if $\overline{A}$ and $\overline{B}$ are vectors of transition matrices, $\overline{A} \otimes \overline{B}$ is also a transition matrix, and if $\overline{A}$ and $\overline{B}$ are vectors of stochastic generators, $\overline{A} \otimes Id_p + Id_n \oplus \overline{B}$ is also a stochastic generator. We already noticed the possible use of classical Kronecker algebra for vector Markov chains whose components are independent. The generalized operations have the same use when the components are not independent. Precisely, if vector $(A^i)_{i \in [1,p]}$ represents the stochastic behavior of process X conditioned on the state of process Y, whose state space is $[1,p]$, this vector is an operand to express the generator matrix of (X,Y).

To generalize these Kronecker operations to more than two operands, the vectors should be essentially of the required dimension, which is stated in the following definition.

Definition 3.3: *If* $(\overline{A}_m)_{m \in [1,c]}$ *is a sequence of vector of matrices such that*
- each component of $\overline{A}_m$ *is a square matrix of dimension* p_m*,*
- the number of components of $\overline{A}_m$ *is* $s_m(\overline{p})$ *with* $\overline{p} = (p_1, \ldots, p_c)$ *and these components are indexed in* $L_m(\overline{p})$.
-the components of $\overline{A}_m$ *are denoted* $(A_m^{\overline{k}})_{\overline{k} \in L_m(\overline{p})}$ *and* $A_m^{\overline{k}} = (a_m^{\overline{k}}(i,j))_{i,j \in [1,p_m]}$

then $C = \bigotimes_{m=1}^{c} \overline{A}_m$ *is defined by:*

$$C = \sum_{\overline{i} \in L(\overline{p})} \bigotimes_{m=1}^{c} l_{i_m}(A_m^{\overline{i}_m}) \qquad or \qquad (\forall \overline{i}, \overline{j} \in L(\overline{p})) \quad c(\overline{i}, \overline{j}) = \prod_{m=1}^{c} a_m^{\overline{i}_m}(i_m, j_m)$$

From simple algebraic arguments, we can show the the Kronecker product of vectors of transition matrices is a transition matrix.

For the sake of completeness, the generalized Kronecker sum is defined as $D = \sum_{m=1}^{c} Id_{r_m(\overline{p})} \otimes \overline{A}_m \otimes Id_{t_m(\overline{p})}$, where $r_m(\overline{p})$ and $t_m(\overline{p})$ have been defined at the beginning of this section. If each $\overline{A}_m$ is a vector of stochastic generators, then D is a stochastic generator.

This last definition concludes the technical introduction on Kronecker products. The important thing is to clarify the relationship between these operators and the identified stochastic behaviors of a vector Markov chain, which is the goal of the following section.

4. Structure of the generator matrix of a vector Markov chain

We resume the discussion of section 2: the first analysis of this chain has lead to the identification of an exhaustive set of dependencies, which are defined to be *probabilistic, concurrent-firing* or *mixed,* and are related to a group of components and a subspace of transitions. The last section gave a short introduction to the tools that will enable us to assemble the building blocks of the global transition rate matrix Q_Z.

We emphasize that the state space E has essentially a structure of cross-product space, and assuming that each space E_i is arbitrarily ordered, the lexicographical ordering on E_Z is used. If $|E_i|$ is finite, E_i is identified to $[1,p_i]$, otherwise to $\mathbb{N}$ and $p_i = \infty$. The matrix $E_{p_i}(x_1,x_2)$ is the matrix of size p_i whose elements in position (y_1,y_2) are defined by $E_{p_i}(x_1,x_2)(y_1,y_2)=1(x_1=y_1)\,1(x_2=y_2)$: E_{p_i} has only one nonzero element, which is on row x_1 and column x_2. At the end of section 2, a partition of the global state space has been defined, and this will be used as a basis for the main theorem. Namely, we recall that

$$E = \bigcup_{\overline{m} \in L(\overline{t})} T_{\overline{m}} \quad \text{and}$$

$$T_{\overline{m}} = \prod_{i=1}^{c} T_i^{m_i} \quad \text{with} \quad \overline{m}=(m_1, \ldots, m_c) \quad \text{and} \quad 1 \le m_i \le t_i$$

Dependencies have been defined in terms of the properties of the transition rates. In this section, the coefficients c_γ, c_λ and c_ϕ, will be structured as elements of vectors of matrices, called *contributions.*

Definition 4.1: *Considering the probabilistic dependency defined by the triple* $(I=\{i\},K,T_I)$, *the contribution of component* X_i *is given by the vector* $(T_I\text{-}\overline{Q}_{X_i \mid Z_K})$, *which has* $\prod_{k \in K} p_k$ *components, and for each value* y_K, *the component* $T_I\text{-}Q_{X_i \mid Z_K=y_K}$ *is a matrix of size* p_i *equal to*

$$(\forall (y_i,y_i') \in E_i \times E_i) \quad such \quad that \quad y_i \ne y_i' \quad (\forall y_K \in E_K)$$

$$T_I\text{-}Q_{X_i \mid Z_K=y_K}(y_i,y_i') = \lim_{t' \to t} \frac{P(X_{i,t'} = y_i' \mid Z_{IUK,t} = y_{IUK})}{t'-t}$$

$$if \quad \{y_i,y_K\} \in proj_{IUK}(T_{I\overline{m}}) \cap R_{IUK}$$

$= 0$ *otherwise*

The diagonal elements of these matrices are chosen so that the sum of each row is zero.

The nonzero coefficients of this vector are exactly the appropriate coefficients $c_\gamma(y, y')$ given in equation 2.1 for a probabilistic dependency γ. When those are not defined, either because $\{y_i, y_K\}$ is not in the designated paving block, or because the conditioning event is not reachable, the value 0 is used.

In a concurrent-firing or mixed dependency, we arbitrarily select a component within the set of moving components: this component is called the *master* of the dependency. For a dependency π defined by $(I_\pi, K_\pi, I_{0_\pi}, p_\pi, f_{i_\pi})$, the master component will bear the responsibility of the coefficient $p(x_{I_\pi}, y_{K_\pi}, y_{I_\pi}')$, which occurs only once for all components in I_π. This choice can be directed by the model itself where this component has a particular leader role. We denote $|f_{I_\pi}|$ the maximum of the cardinalities of the sets $f_{I_\pi}(y_{K_\pi})$: $|f_{I_\pi}| = \max_{\{x_{I_\pi}, y_{K_\pi}\} \in R_{I_\pi U K_\pi}} (|f_{I_\pi}(y_{K_\pi})|)$. The number $|f_{I_\pi}|$ represents the maximum number of possible destinations for this dependency. Moreover, we assume that each of the sets $f_{I_\pi}(y_{K_\pi})$ is ordered by the trace of the lexicographical ordering, and we denote accordingly: $f_{I_\pi}(y_{K_\pi}) = \{x_{I_\pi, l}'\}_{l \in [1, |f_{I_\pi}(y_{K_\pi})|]}$.

Definition 4.2: *Considering the concurrent-firing dependency defined by the 5-tuple* $(I, K, T_I = \{x_I\}, p, f_I)$, *for all* $i \in I$, *the contribution of* X_i *is given by a sequence of* $|f_I| + 1$ *vectors* $(T_I - \overline{Q}_{X_i \mid Z_K}(l))_{l \in [0, |f_I|]}$ *which have* $\prod_{k \in K} p_k$ *components and*

(i) if i *is not the master and for each value* $\{x_I, y_K\} \in R_{KUI}$, *the component* $T_I - Q_{X_i \mid Z_K = y_K}(0)$ *is the matrix* $E_{p_i}(x_i, x_i)$, *the components* $(T_I - Q_{Z_I \mid Z_K = y_K}(l))_{l \in [1, |f_I(y_K)|]}$ *are respectively the matrices* $E_{p_i}(x_i, proj_i(x_{I,l}'))$, *and for* $l \in]\,|f_I(y_K)|, |f_I|\,]$ *the components are null. If* $\{x_I, y_K\} \notin R_{KUI}$ *all vectors have null components.*

(ii) if i *is the master, and for each value* $\{x_I, y_K\} \in R_{KUI}$, *the component* $T_I - Q_{X_i \mid Z_K = y_K}(0)$ *is the matrix* $(-\sum_{x_{I,l}' \in f_I(y_K)} p(x_I, y_K, x_{I,l}')) \cdot E_{p_i}(x_i, x_i)$, *the components* $(T_I - Q_{X_i \mid Z_K = y_K}(l))_{l \in [1, |f_I(y_K)|]}$ *are respectively the matrices* $p(x_I, y_K, x_{I,l}') \cdot E_{p_i}(x_i, proj_i(x_{I,l}'))$, *and for* $l \in]\,|f_I(y_K)|, |f_I|\,]$ *the components are*

null. If $\{x_I, y_K\} \notin R_{KUI}$*, the vectors have uniformly null components.*

This definition requires some comments. A contribution is given knowing the value y_k of the interacting component and the starting state x_I of the moving components, which is fixed by definition. Depending on the value of y_K, the number of possible destinations of this concurrent-firing dependency may vary. This is why we have a number of contributions equal to $|f_I|$, the maximum number of destinations. Whenever a given value y_K has fewer possible destinations, extra vectors are filled with null values. We must notice that the contributions of a master component differ from the other ones only by a scalar multiple. For a concurrent-firing dependency λ, the coefficient c_λ of equation 2.3 will be obtained as an element of a Kronecker product of these matrices.

We define next the contributions of a mixed dependency where we will be able to recognize both the probabilistic and concurrent firing patterns.

Definition 4.3: *Considering the mixed dependency defined by the 6-tuple* $(I, I_0, K, T_I=\{x_I\}, p, f_I)$*, for all* $i \in I$*, the contribution of* X_i *is given by a sequence of* $|f_I|+1$ *vectors* $(T_I\text{-}\overline{Q}_{X_i \mid Z_K}(l))_{l \in [0, |f_I|]}$ *and one vector* $(T_I\text{-}\overline{Q}_{X_i \mid Z_K})$ *for each* $i \in I_0$*, which all have* $\prod_{k \in K} p_k$ *components and*

(i) if i *is not the master and for each value* $\{x_I, y_K\} \in R_{KUI}$*, the component* $T_I\text{-}Q_{X_i \mid Z_K = y_K}(0)$ *is the matrix* $E_{p_i}(x_i, x_i)$*, the components* $(T_I\text{-}Q_{Z_I \mid Z_K = y_K}(l))_{l \in [1, |f_I(y_K)|]}$ *are respectively the matrices* $E_{p_i}(x_i, proj_i(x_{I,l}'))$*, and for* $l \in]\,|f_I(y_K)|, |f_I|\,]$ *the components are null. If* $\{x_I, y_K\} \notin R_{KUI}$ *the vectors have null components.*

(ii) if i *is the master and for each value* $\{x_I, y_K\} \in R_{KUI}$*, the component* $T_I\text{-}Q_{X_i \mid Z_K = y_K}(0)$ *is the matrix* $(-\sum_{x_{I,l}' \in f_I(y_K)} p(x_I, y_K, x_{I,l}')) \,.\, E_{p_i}(x_i, x_i)$*, the components* $(T_I\text{-}Q_{X_i \mid Z_K = y_K}(l))_{l \in [1, |f_I(y_K)|]}$ *are respectively the matrices* $p(x_I, y_K, x_{I,l}') \,.\, E_{p_i}(x_i, proj_i(x_{I,l}'))$*, and for* $l \in]\,|f_I(y_K)|, |f_I|\,]$ *the components are null. If* $\{x_I, y_K\} \notin R_{KUI}$ *the vectors have uniformly null components.*

(iii) for $i \in I_0$*, the vector* $(T_I\text{-}\overline{Q}_{X_i \mid Z_K})$ *has components of size* p_i *equal to*

$$(\forall (y_i, y_i') \in E_i \times E_i) \quad \text{such that} \quad y_i \neq y_i' \quad (\forall y_K \in E_K)$$

$$T_I\text{-}Q_{X_t \mid Z_K=y_K}(y_i, y_i') = \lim_{t' \to t} \frac{P(X_{i,t'} = y_i' \mid Z_{IUK,t} = y_{IUK})}{t' - t}$$

$$\text{if} \quad \{x_i, y_K\} \in proj_{IUK}(T_{\overline{m}}) \cap R_{IUK}$$

$$= 0 \qquad \text{otherwise}$$

These matrices are completed with diagonal elements to enforce the sum of each row to be zero.

Cases (i) and (ii) of this definition are in all points identical to those of definition 4.2 and case (iii) is the replication of definition 4.1 for all components in I_0. Generally any compound structure should require contributions from all of its elementary trends. The following definition will modify these contribution vectors to make them usable in a generalized Kronecker product.

Definition 4.4: *In the framework of a problem in* $E = E_1 \times \cdots \times E_c$, *denoting* $I = \{i\}$ *and given a vector of matrices of the type* $\overline{Q}_{X_t \mid Z_K}$ *that has* $\prod_{k \in K} p_k$ *components in* $M(p_i)$, *the expanded version of this vector is* $e\left[\overline{Q}_{X_t \mid Z_K}\right] = (Q_{X_t \mid Z_{\overline{I}} = x_{\overline{I}}})_{x_{\overline{I}} \in E_{\overline{I}}}$ *that has* $\prod_{k \neq i} p_k$ *components in* $M(p_i)$ *such that*

$$Q_{X_t \mid Z_{\overline{I}} = x_{\overline{I}}} = Q_{X_t \mid Z_K = x_K} \qquad \forall x_{\overline{I}} \in E_{\overline{I}} \cap R_{\overline{I}} \qquad \text{with} \qquad x_K = proj_K(x_{\overline{I}})$$

$$= 0 \qquad \text{otherwise}$$

The vectors of matrices defined as *contributions* in the preceding definitions lead to inhomogeneous vector sizes, which are based on the set of interacting components K of each dependency. The expansion is a mapping that gives a regular size to each vector, allowing the use of definition 3.3 in the following theorem.

These contributions will be embedded into the general formula within noticeable terms. Here also we emphasize that two different constructions arise for probabilistic and concurrent-firing elements.

Definition 4.5: *Considering vectors of contributions as defined previously, we want to define their related term:*

(i) If the contribution $(T_I\text{-}\overline{Q}_{X_t \mid Z_K})$ *comes from a probabilistic dependency, the related term is defined by*

$$t\left[T_I - \overline{Q}_{X_i \mid Z_K}\right] = \otimes_{i=k}^{c} \overline{M}_k$$

where $\overline{M}_i$ *is equal to* $e\left[T_I - \overline{Q}_{X_i \mid Z_K}\right]$. *For* $k \neq i$, $\overline{M}_k$ *has all its components equal to* Id_{p_k}.

(ii) If this contribution $T_I \text{-} \overline{Q}_{X_i \mid Z_K}(l)$ *comes from a concurrent-firing dependency involving the moving components* Z_I, *the related term is defined relatively to* I *globally, and* l *by:*

$$t\left[T_I - \overline{Q}_{X_i \mid Z_K}(l)\ ,\ i \in I\right] = \otimes_{i=k}^{c} \overline{M}_k$$

where $\overline{M}_k$ *is equal to* $e\left[T_I \text{-} \overline{Q}_{X_k \mid Z_K}(l)\right]$, *if* $k \in I$. *For* $k \notin I$, $\overline{M}_k$ *has all its components equal to* Id_{p_k}.

The common feature of these terms is that all moving components introduce specific contributions as the nonmoving components show their neutrality with a participation equal to an identity matrix. Mixed dependencies have both patterns (i) and (ii).

Now we proceed to the main theorem of this section. We still consider that from an exhaustive analysis of the chain, we withdraw $\Pi\Gamma$ the set of all probabilistic dependencies and $\Pi\Lambda$ the set of all concurrent-firing dependencies and $\Pi\Phi$ the set of all mixed dependencies.

Theorem 4.1: *Given an exhaustive analysis, the generator matrix of the global model is such that:*

$$(\forall y \in R\ )\ \ (\forall y' \in E\ )$$

$$Q_Z(y, y') = \sum_{\pi \in \Pi\Gamma} T_{I_\pi} \text{-} Q_Z(y, y') + \sum_{\pi \in \Pi\Lambda} T_{I_\pi} \text{-} Q_Z(y, y') + \sum_{\pi \in \Pi\Phi} T_{I_\pi} \text{-} Q_Z(y, y') \quad 4.0$$

-if π *is a probabilistic dependency, then*

$$T_{I_\pi} \text{-} Q_Z = t\left[T_{I_\pi} \text{-} \overline{Q}_{X_{i_\pi} \mid Z_{K_\pi}}\right] \quad 4.1$$

-if π *is a concurrent-firing dependency, then*

$$T_{I_\pi} \text{-} Q_Z = \sum_{l \in [0, \mid f_{I_\pi} \mid]} t\left[T_{I_\pi} \text{-} \overline{Q}_{X_i \mid Z_{K_\pi}}(l)\ ,\ i \in I_\pi\right] \quad 4.2$$

-if π *is a mixed dependency, then*

$$T_{I_\pi} \text{-} Q_Z = \sum_{l \in [0, \mid f_{I_\pi} \mid]} t\left[T_{I_\pi} \text{-} \overline{Q}_{X_i \mid Z_{K_\pi}}(l)\ ,\ i \in I_\pi\right] + \sum_{i \in I_0} t\left[T_{I_\pi} \text{-} \overline{Q}_{X_i \mid Z_{K_\pi}}\right] \quad 4.3$$

Proof: This proof will show that the result of this matrix expression is identical to the result given by Theorem 2.1. We denote A the matrix

$$A = \sum_{\pi \in \Pi\Gamma} T_{I_\pi}\text{-}Q_Z + \sum_{\pi \in \Pi\Lambda} T_{I_\pi}\text{-}Q_Z + \sum_{\pi \in \Pi\Phi} T_{I_\pi}\text{-}Q_Z$$

To analyze this formula, we restrict y in any one of the paving blocks, namely $y \in T_{\overline{m}} \cap R$ and $y' \in E$, $y \neq y'$. We have four subsets Γ, Λ, Φ and Φ' related to $T_{\overline{m}}$, which are, respectively, the list of probabilistic dependencies, the list of consistent concurrent-firing dependencies, the lists of consistent mixed dependencies and partially consistent mixed dependencies. From there, we define the sets of moving components $(I_\gamma)_{\gamma\in\Gamma}$, $(I_\lambda)_{\lambda\in\Lambda}$, $(I_\phi)_{\phi\in\Phi}$, $(J_{\phi'})_{\phi'\in\Phi'}$ (where J is the trace related to $T_{\overline{m}}$ as stated in definition 2.5), which altogether form a partition of $[1,c]$. We study the element (y, y') of these matrices:

(a) For all $\gamma \in \Pi\Gamma$ and $y \neq y'$ we have from 4.1

$$T_{I_\gamma}\text{-}Q_Z(y,y') = \left(\prod_{k=1, k\neq i_\gamma}^{c} Id_{p_k}(y_k, y_k') \right) \times T_{i_\gamma}\text{-}Q_{X_{i_\gamma} \mid Z_{K_\gamma} = y_{K_\gamma}}(y_{i_\gamma}, y_{i_\gamma}') \qquad 4.4$$

If γ is inconsistent with $T_{\overline{m}}$, then $proj_{i_\gamma}(y) \notin proj_{i_\gamma}(T_{\overline{m}})$. This implies that $T_{i_\gamma}\text{-}Q_{X_{i_\gamma} \mid Z_{K_\gamma}=y_{K_\gamma}}(y_{i_\gamma}, y_{i_\gamma}') = 0$, leading to a zero result for the product. On the other hand, if γ is consistent with $T_{\overline{m}}$

$$\begin{aligned} T_{I_\gamma}\text{-}Q_Z(y,y') &= \lim_{t' \to t} \frac{P(X_{i_\gamma t'} = y_{i_\gamma}' \mid Z_{I_\gamma \cup K_\gamma t} = y_{I_\gamma \cup K_\gamma})}{t' - t} \times 1(y_{\overline{I}_\gamma} = y_{\overline{I}_\gamma}') \\ &= c_\gamma(y, y') \end{aligned}$$

(b) For all $\lambda \in \Pi\Lambda$ and $y \neq y'$ we have from 4.2

$$T_{I_\lambda}\text{-}Q_Z(y,y') = \qquad 4.5$$

$$\left(\prod_{k=1, k\notin I_\lambda}^{c} Id_{p_k}(y_k, y_k') \right) \times \Big(- \sum_{l \in [1, \mid f_{I_\lambda}(y_{K_\lambda}) \mid]} p(x_{I_\lambda}, y_{K_\pi}, x_{I_\lambda, l}') \times \left(\prod_{k \in I_\lambda} E_{p_k}(x_k, x_k)(y_k, y_k') \right)$$

$$+ \sum_{l \in [1, \mid f_{I_\lambda}(y_{k_\lambda}) \mid]} \left(\prod_{k=1, k\notin I_\lambda}^{c} Id_{p_k}(y_k, y_k') \right) \times p(x_{I_\lambda}, y_{K_\lambda}, x_{I_\lambda, l}') \times \left(\prod_{k \in I_\lambda} E_{p_k}(x_k, proj_k(x_{I_\lambda, l}'))(y_k, y_k') \right)$$

If λ is inconsistent with $T_{\overline{m}}$, there exist $k_0 \in I_\lambda$ such that $proj_{k_0}(y) \notin proj_{k_0}(T_{\overline{m}}) = \{x_{k_0}\}$. This implies that the contribution of X_{k_0} is such that $E_{p_{k_0}}(x_{k_0}, x_{k_0})(y_{k_0}, y_{k_0}') = E_{p_{k_0}}(x_{k_0}, proj_{k_0}(x_{I_\phi, l}'))(y_{k_0}, y_{k_0}') = 0$, leading to a zero result for the product.

On the other hand if λ is consistent with $T_{\overline{m}}$

$$T_{I_\lambda} - Q_Z(y, y') = \sum_{x_{I_\lambda}' \in f_{I_\lambda}(y_{K_\lambda})} p(x_{I_\lambda}, y_{K_\lambda}, x_{I_\lambda}') \times 1(y_{I_\lambda}' = x_{I_\lambda}') \times 1(y_{\overline{I}_\lambda} = y_{\overline{I}_\lambda}')$$
$$= c_\lambda(y, y')$$

Under the condition $y \neq y'$, the first term of the sum 4.5 is always null. This term is a diagonal matrix, which allows fulfillment of the normalizing condition as we will see further.

(c) If ϕ a mixed dependency and $y \neq y'$, we have from 4.3

$$T_{I_\phi} - Q_Z(y, y') = \tag{4.6}$$
$$\left(\prod_{k=1, k \notin I_\phi}^{c} Id_{p_k}(y_k, y_k') \right) \times \left(- \sum_{l \in [1, |f_{I_\phi}(y_{K_\phi})|]} p(x_{I_\phi}, y_{K_\phi}, x_{I_\phi, l}') \times \left(\prod_{k \in I_\phi} E_{p_k}(x_k, x_k)(y_k, y_k') \right) \right.$$
$$+ \sum_{l \in [1, |f_{I_\phi}(y_{K_\phi})|]} \left(\prod_{k=1, k \notin I_\phi}^{c} Id_{p_k}(y_k, y_k') \right) \times p(x_{I_\phi}, y_{K_\phi}, x_{I_\phi, l}') \times \left(\prod_{k \in I_\phi} E_{p_k}(x_k, proj_k(x_{I_\phi, l}'))(y_k, y_k') \right)$$
$$+ \sum_{i \in I_{0_\phi}} \left(\prod_{k=1, k \neq i}^{c} Id_{p_k}(y_k, y_k') \right) \times T_i - Q_{X_i \mid Z_{K_\phi} = y_{K_\phi}}(y_i, y_i')$$

If ϕ is totally unconsistent with $T_{\overline{m}}$ or partially consistent but its trace J_ϕ is such that $J_\phi \cap I_{0_\phi} = \emptyset$, then the arguments given in (a) and (b) show that $T_{I_\phi} - Q_Z(y, y')$ is null. If ϕ is partially inconsistent, but $J_\phi \cap I_{0_\phi} \neq \emptyset$, still with identical arguments, we show that only the part $\sum_{i \in I_{0_\phi} \cap J_\phi} \left(\prod_{k=1, k \neq i}^{c} Id_{p_k}(y_k, y_k') \right) \times T_{i_\phi} - Q_{X_{i_\phi} \mid Z_{K_\phi} = y_{K_\phi}}(y_{i_\phi}, y_{i_\phi}')$ leads to nonzero result. Namely, if $\phi \in \Phi$ is inconsistent with $T_{\overline{m}}$ but $J_\phi \cap I_{0_\phi} \neq \emptyset$

$$T_{I_\phi} - Q_Z(y, y') = \sum_{i \in I_{0_\Phi} \cap J_\Phi} \lim_{t' \to t} \frac{P(X_{i, t'} = y_i \mid Z_{IUK, t} = y_{IUK})}{t' - t} 1(y_i \neq y_i') \times \prod_{k \neq j} 1(y_k = y_k')$$
$$= c_\phi(y, y')$$

If ϕ is consistent with $T_{\overline{m}}$ then

$$T_{I_\phi} - Q_Z(y, y') = \sum_{x_{I_\phi}' \in f_{I_\phi}(y_{K_\phi})} p(x_{I_\phi}, y_{K_\phi}, y_{I_\phi}') \times 1(y_{I_\phi}' = x_{I_\lambda}') \times 1(y_{\overline{I}_\phi} = y_{\overline{I}_\phi}')$$
$$+ \sum_{i \in I_0} \lim_{t' \to t} \frac{P(X_{i, t'} = y_i \mid Z_{IUK, t} = y_{IUK})}{t' - t} 1(y_i \neq y_i') \times \prod_{k \neq j} 1(y_k = y_k')$$
$$= c_\phi(y, y')$$

as defined in 2.3(i).

Summarizing the points (a),(b) and (c), we have shown that

$$(\forall y \in T_{\overline{m}} \cap R)\ (\forall y' \in E) \qquad y \neq y'$$

$$A(y,y') = \sum_{\gamma\in\Gamma} c_\gamma(y,y') + \sum_{\lambda\in\Lambda} c_\lambda(y,y') + \sum_{\phi\in\Phi} c_\phi(y,y') + \sum_{\phi'\in\Phi'} c_{\phi'}(y,y')$$

which shows the equality $A = Q_Z$ under the preceding conditions, if we compare this result with the statement of Theorem 2.1. This is true for any $T_{\overline{m}}$.

To conclude the proof, we must show the equality for the diagonal terms in R: the idea is to show that on both sides of equality 4.0, we have the row normalization property. The matrix Q_Z clearly has this property from its stochastic definition. Moreover, this property is stable by matrix addition, and we are going to prove it for each T_{I_π}-Q_Z matrix:

(a) if γ is a probabilistic dependency, and for all $y, y' \in E$, the terms on row y are

$$\left(\prod_{k=1,k\neq i_\gamma}^{c} Id_{p_k}(y_k, y_k')\right) \times T_{i_\gamma}\text{-}Q_{X_{i_\gamma} | Z_{K_\gamma}=y_{K_\gamma}}(y_{i_\gamma}, y_{i_\gamma}')$$

which clearly sum to 0, as T_{i_γ}-$Q_{X_{i_\gamma} | Z_{K_\gamma}=y_{K_\gamma}}$ does.

(b) If λ is a concurrent-firing dependency, for all $(y, y') \in R \times E$, row y of T_{I_π}-Q_Z is

$$\left(\prod_{k=1,k\notin I_\lambda}^{c} Id_{p_k}(y_k, y_k')\right) \times \left(-\sum_{l\in[1,|f_{I_\lambda}(y_{K_\lambda})|]} p(x_{I_\lambda}, y_{K_\pi}, x_{I_\lambda,l}') \times \left(\prod_{k\in I_\lambda} E_{p_k}(x_k, x_k)(y_k, y_k')\right)\right.$$

$$+ \sum_{l\in[1,|f_{I_\lambda}(y_{k_\lambda})|]} \left(\prod_{k=1,k\notin I_\lambda}^{c} Id_{p_k}(y_k, y_k')\right) \times p(x_{I_\lambda}, y_{K_\lambda}, x_{I_\lambda,l}') \times \left(\prod_{k\in I_\lambda} E_{p_k}(x_k, proj_k(x_{I_\lambda,l}'))(y_k, y_k')\right)$$

The only row with nonzero terms are those corresponding to a state y such that $proj_{I_\pi}(y) = x_{I_\pi}$. On these rows, the terms are respectively $\left(p(x_{I_\lambda}, y_{K_\lambda}, x_{I_\lambda,l}')\right)_{l\in[1,f_{I_\pi}(y_k)]}$ and $-\sum_{l\in[1,f_{I_\pi}(y_k)]} p(x_{I_\lambda}, y_{K_\lambda}, x_{I_\lambda,l}')$, which sum up to 0.

(c) If π is a mixed dependency then both terms (a) and (b) appear and the stability by addition leads the result.

This argument concludes the proof.

We have discarded in both theorems the transitions defined by (y, y'), such that y is not reachable. We know that if $y \in T_{\overline{m}}$, only dependencies Γ, Λ, Φ and Φ', as defined previously, lead to a nonzero result for the corresponding $T_{I_\pi} - Q_Z$. If we have the property

$$(\forall \pi \in \Gamma \cup \Lambda \cup \Phi \cup \Phi') \quad y_{K_\pi U I_\pi} \notin R_{K_\pi U I_\pi}$$

$A(y, y')$ will be equal to zero. This comes from the fact that contributing values were set to zero. But there might be models where this last property does not hold. Element y can be non-reachable but all the $y_{K_\pi U I_\pi}$ can be reachable projections. The corresponding $A(y, y')$ element will then depend on the particular model.

That problem will be addressed by choosing properly the numerical resolution method: If we consider using a power method [Stew80] to compute the steady state vector of this chain, then

- the Kronecker algebra is of an essential use to reduce the computation and storage cost [Davi81]. As a matter of fact, there is no need at all to expand matrix A, which is usually huge, to compute the regular product $V\ A$. This is done by replacing the large matrix product by a sequence of smaller products.

- For this power method, we need to know the reachability space R, and the initial vector V_0 must be chosen such that

$$(\forall y' \notin R) \quad V_0(y') = 0$$

We know that A has the same property as Q, namely:

$$(\forall y \in R) \text{ and } (\forall y' \in E - R) \quad A(y, y') = 0$$

so that $V_1 = V_0 A$ is defined by

$$V_1(z) = \sum_{y \in E} V_0(y) A(y, z) = \sum_{y \in R} V_0(y) A(y, z)$$

Whenever $z \in E - R$, this gives $V_1(z) = 0$, yielding that $V_1(z)$ is also in R. If A_R and $V_{0,R}$ denote the restrictions of the preceding A and V_0 to R, this shows that it is equivalent to study either one of the sequences $V_0.A^n$ or $V_{0,R}.A_R^n$, $n \in \mathbb{N}$, knowing that A_R has the nice property of irreducibility. The recurrence that is used is the following: Let σ the maximum of the absolute value of the diagonal elements of matrix A, and ϵ be a small positive number. Then given $W_0 \in R$, the sequence

$$W_n = W_{n-1} + \frac{1}{\sigma(1+\epsilon)} W_{n-1} A$$

is assured to converge.

These are the algorithmic advantages one can draw from the formula. Nevertheless the power method applied to this type of matrix can probably be refined. This is the subject of continuing work.

5. Conclusion

To summarize what we have discussed so far, we may say:

- We have decomposed the transition rates of matrix Q_Z by using the properties of the stochastic behavior of the chain.
- We have given two building rules, for probabilistic and concurrent-firing dependencies based on elementary matrices and Kronecker algebra.
- The main result shows how these match on a problem that fits into the framework we have defined, for all relevant transitions.

Our interest in this technique lies in the belief that it will enable us to study the behavior of the chain Z by looking at small groups of components over subspaces of transitions where their stochastic properties can be clearly identified. The probabilistic dependency introduces loose constraints among components (a particular case being pure independence), as the concurrent-firing and mixed dependencies bring more constraints. A reasonable decomposition is achievable only if the problem has few of these last singular points that we called concurrent-firing dependencies or if they show a regular pattern. Our feeling is that this situation might be common when modeling distributed applications where, it is reasonable to assume that, when a message is sent, its transmission time is independent of the forthcoming processors activities. Moreover, we think that Kronecker algebra is a powerful tool to express vectorial problems that show actual dependencies among components. Indeed, we have been able to derive a closed form expression of the generator of a problem without any size limitations. Moreover, the numerical procedures are efficient as they use the structural prooerties of the generator.

To conclude this paper, we will emphasize the relationships with other related work. The first related area is the field of stochastic Petri nets. This technique also uses a graphic representation of the system under study and an underlying Markov model [Moll82], [DTGN84], [MaBC84], [VeHo86]. After the description phase, a program is run to derive the reachability graph, to generate the stochastic matrix and to compute the steady state probabilities and the performance. The line of thought of the present method is identical, except that it attempts to push further the analytical study and to exploit the result in the numerical algorithmic part. It would be interesting to see whether a similar approach can be pursued using a Petri net specification of the problem.

Another related area is the well-known field of queuing theory. Previous work has been done [Neut81], [Mass84], to derive analytic results or bounds of single queues or network of queues using Kronecker algebra. A systematic method to derive bounds within the framework presented here would be of great interest.

References

[CaRo83]

O. Carvalho, G. Roucairol: "On mutual exclusion in computer network", *Communications of the ACM, Vol 26, No 2, Feb 1983.*

[ChMi81]

K. Chandy, J. Misra: "Asynchronous distributed simulation via a sequence of parallel computations", *Communications of the ACM, Vol 24, No 4, April 1981.*

[ChLa85]

K.M. Chandy, L. Lamport: "Distributed snapshots: Determining global states of distributed systems", *ACM Transactions on Computer Systems, Vol 3, No 1, Jan 1985, pp 63-85.*

[DFVG83]

E.W. Dijkstra, W.H.J. Feijen, A.J.M. Van Gasteren: "Derivation of a termination detection algorithm for distributed computations", *I.P.L., Vol 16, June 1983.*

[Davi81]

M. Davio: "Kronecker products and shuffle algebra", *IEEE Transactions on Computers, Vol C-30, No 2, Feb 81.*

[DTGN]

J.B. Dugan, K.S. Trivedi, R.M. Geist, V.F. Nicola: "Extended stochastic Petri nets: Applications and analysis", *Performance'84, pp 507-519, Paris, France, December 1984.*

[Lamp78]

L. Lamport: "Times, clocks and events and the ordering of events in a distributed system", *Communications of the ACM, Vol 21, No 7, July 1978.*

[MaBC84]

M.A. Marsan, G. Balbo, G. Conte: "A class of generalized stochastic Petri nets", *ACM Transactions on Computer Systems, Vol 2, pp 93-122, May 1984.*

[Mass84]

W.A. Massey: "Open networks of queues: Their algebraic structure and estimating their transient behavior", *Advances in Applied Probability, Vol 16, No 1, March 1984.*

[Maza83]

N. Mazars: "Description des systemes de dependances stochastiques en theorie de la fiabilite", *Cahier du bureau universitaire de la recherche operationnelle, No 39,40, Paris, 1983*

[MeMu79]

D.A. Menasce, R.R. Muntz: "Locking and deadlock detection in a distributed data-base", *I.E.E.E. Transactions on S.E., Vol SE 5,3, May 1979.*

[Moll82]

M. Molloy: "Performance analysis using stochastic Petri nets", *IEEE Transactions on Computers, Vol C-31, No 9, Sept 1982.*

[Neut81]

M.F. Neuts: "Matrix geometric solutions in stochastic models, an algorithmic approach", *The John Hopkins University Press, 1981.*

[Plat84]

B.D. Plateau: "De l'evaluation du parallelisme et de la synchronisation", *These d'etat, ISEM, Universite de Paris-Sud, 91405 Orsay, France, Nov 1985.*

[Plat85]

B.D. Plateau: "On the stochastic structure of parallelism and synchronization models for distributed algorithms", *ACM Sigmetrics Conference on Measurement and Modelling of Computer systems", Austin, August 1985.*

[PlTr86]

B.D. Plateau, S. Tripathi: "Performance analysis of synchronization for two communicating processes", *Internal report ISEM, No 23, Universite de Paris-Sud, 91405 Orsay, France, Feb. 1985.*

[Rayn85]

M. Raynal: "Algorithmes distribues et protocoles", *Eyrolles, Paris, 1985.*

[RiAg81]

G. Ricart, A. Agrawala: "An optimal algorithm for mutual exclusion in computer network", *Communications of the ACM, Vol 24, No 1, Jan 81.*

[Stew80]

W. Stewart, E. Gelenbe, J. Labetoulle, M. Metivier, G. Pujolle: "Reseaux de files d'attente, modelisation et traitement numerique", *Ed.des hommes et techniques, Monographies informatiques de l'AFCET, Chap 4, 1981.*

[VeHo86]

M. Vernon, M. Holliday: "Performance analysis of multiprocessor cache consistency protocols using generalized timed Petri nets", *Proc. of Performance'86 and ACM Sigmetrics 1986, May 27-30, Raleigh, pp 9-17.*

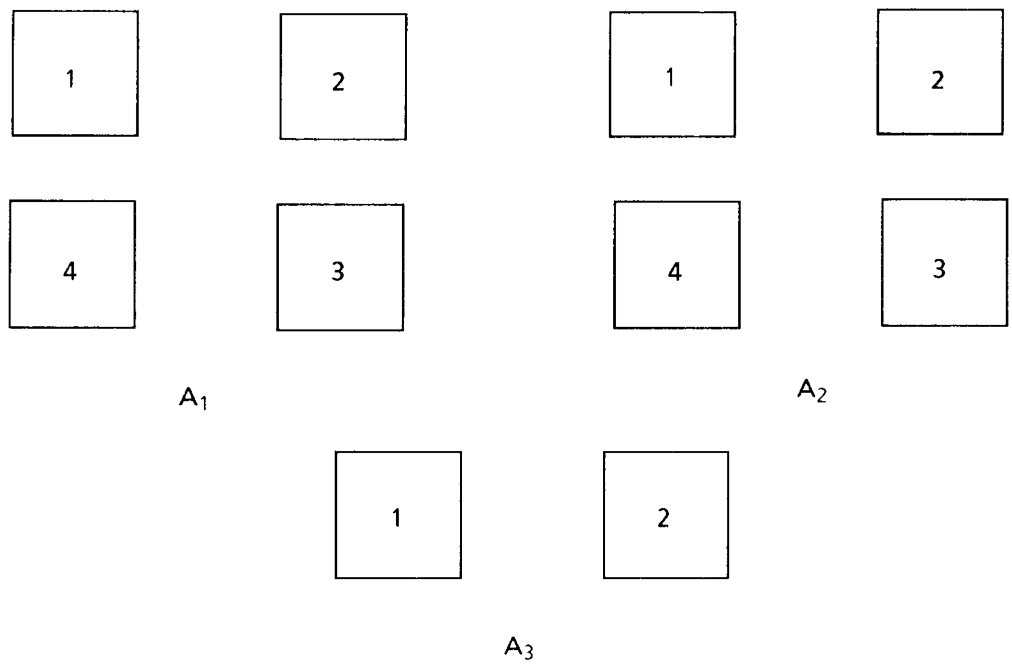


Figure 1

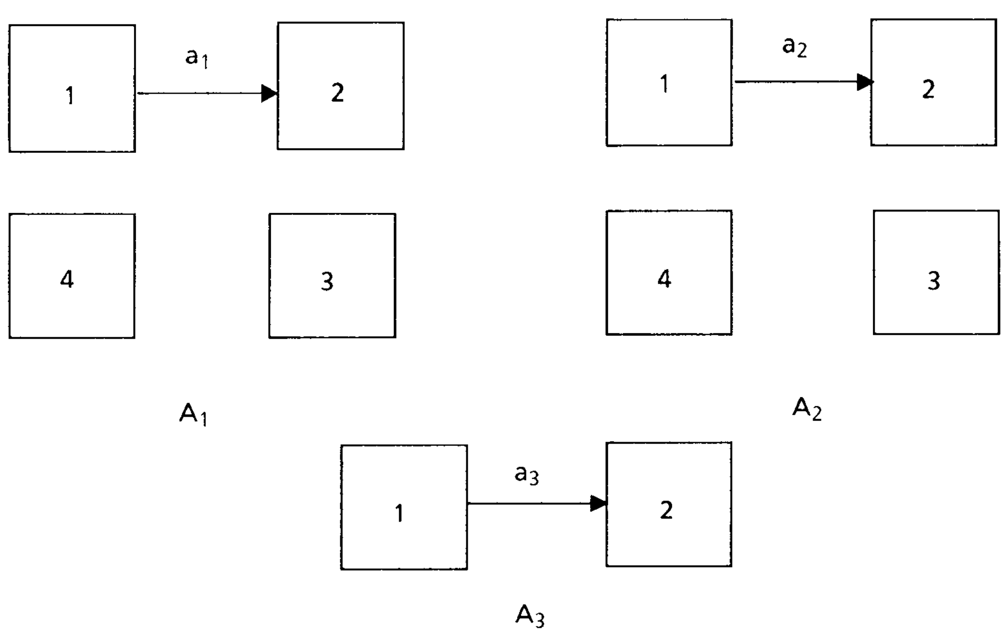


Figure 2

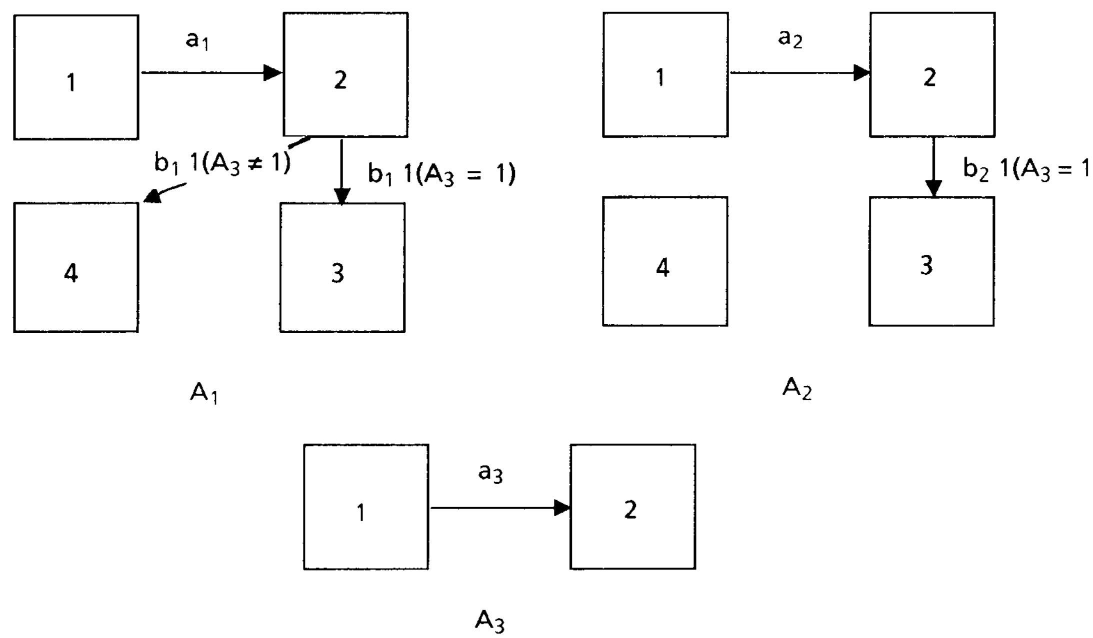


Figure 3

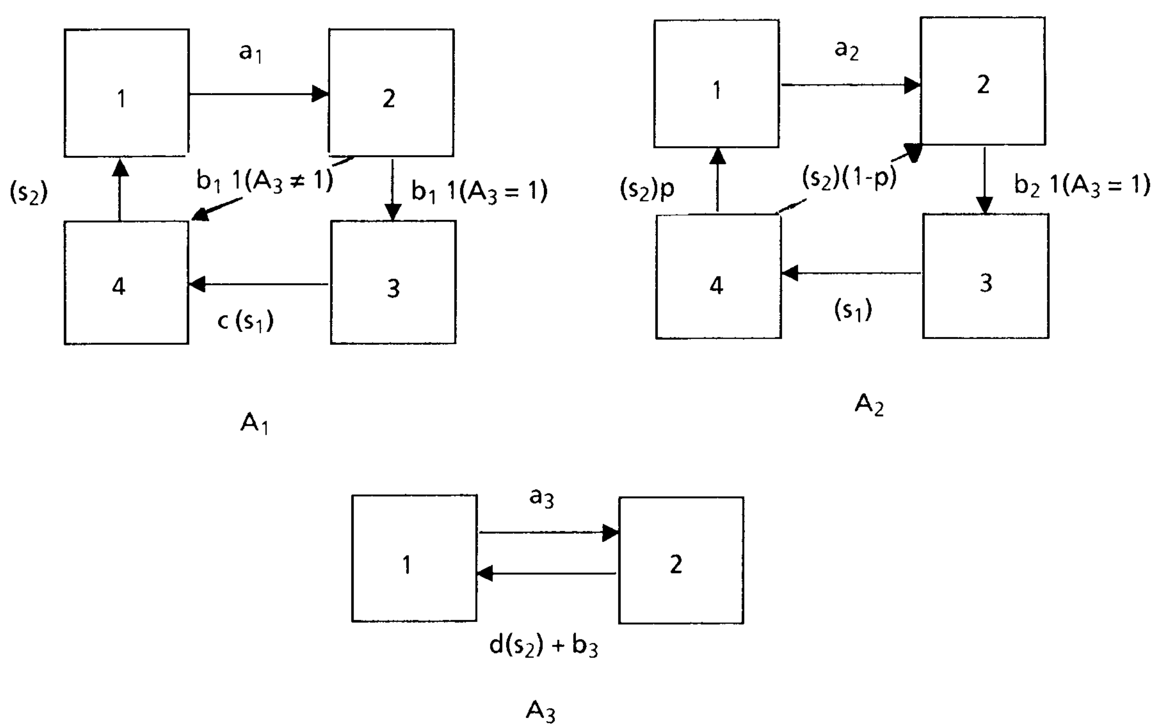
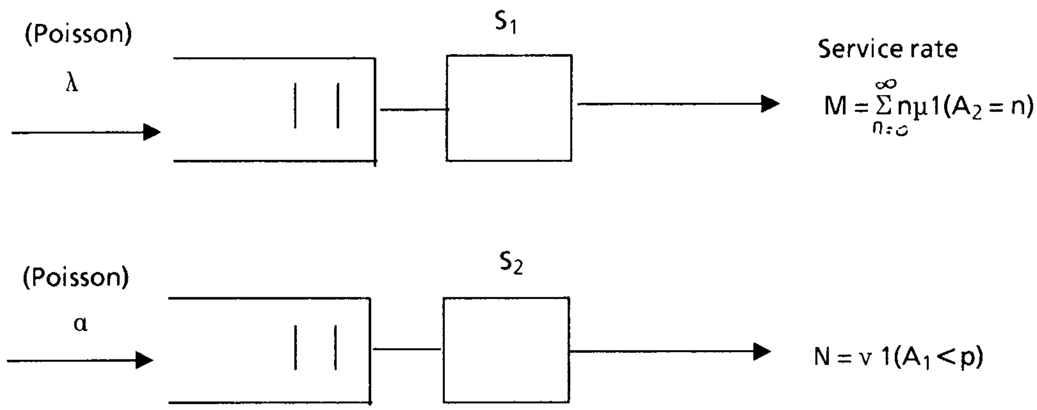


Figure 4



A_1 and A_2 represent respectively the number of customers in servers 1 and 2 at the current time. The service rates of S_1 and S_2 are respectively μ and v

The state transition diagrams are:

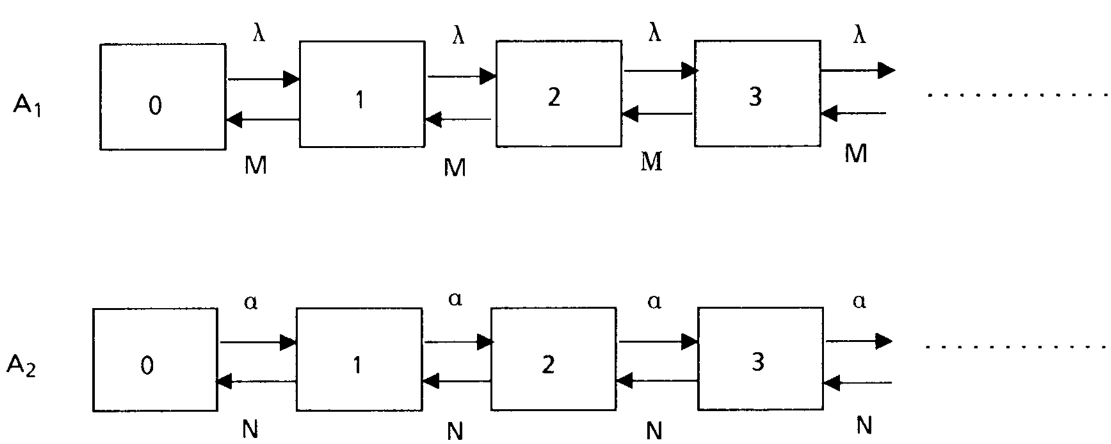
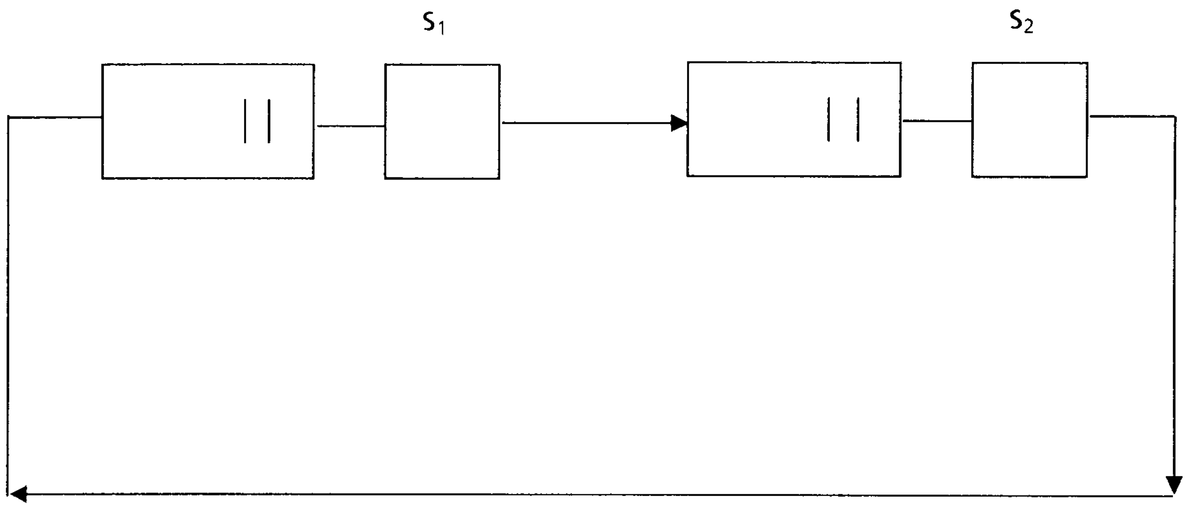


Figure 5



The service rate of S_1 is μ and the service rate of S_2 is v. The total number of customers is N.

Graphs A_1 and A_2 represents respectively the number of customers in S_1 and S_2. Their transitions are:

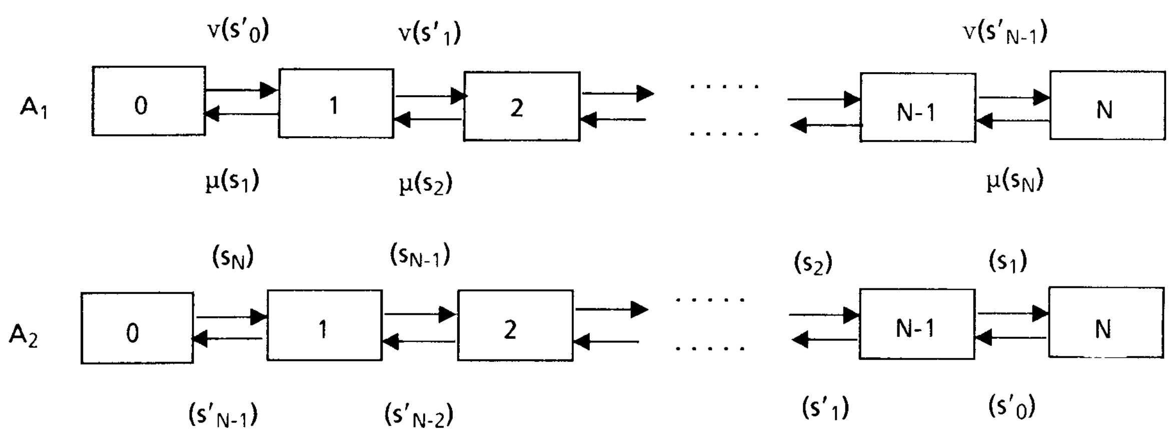


Figure 6

APPENDIX

To give an idea of the applicability of the method, we derive the stochastic generator of the Markov chain $Z_t = (X_{1,t}, X_{2,t}, X_{3,t})$, whose behavior is described by the stochastic graph network (A_1, A_2, A_3) of Figure 4. The results of the analysis are summarized in the following tables. Matrices $(M_i)_{i=1,6}$ are given at the end. The first table is derived from Figure 3.

Probabilistic Dependencies		
Description	Contribution	Related term
$I=\{1\}$, $K=\{3\}$, $T=[1,2]$	$\overline{Q}_{X_1 \mid X_3} = (M_1, M_2)$	$e\left[\overline{Q}_{X_1 \mid X_3}\right] \otimes Id_4 \otimes Id_2$
$I=\{2\}$, $K=\{3\}$, $T=[1,2]$	$\overline{Q}_{X_2 \mid X_3} = (M_3, M_4)$	$Id_4 \otimes e\left[\overline{Q}_{X_2 \mid X_3}\right] \otimes Id_2$
$I=\{3\}$, $K=\emptyset$, $T=[1]$	$Q_{X_3} = (M_5)$	$Id_4 \otimes Id_4 \otimes e\left[Q_{X_3}\right]$

The next table describes the dependency labeled (s_1).

Concurrent-firing Dependencies		
Description	Contribution	Related term
$I=\{1,2\}$, $K=\emptyset$, $T=[3]\times[3]$ Nb of destinations $\mid f \mid = 1$	for X_1: $Q_{X_1}(0) = -c\ E_4(3,3)$ $Q_{X_1}(1) = c\ E_4(3,4)$ for X_2: $Q_{X_2}(0) = E_4(3,3)$ $Q_{X_2}(1) = E_4(3,4)$	$-c\ E_4(3,3) \otimes E_4(3,3) \otimes Id_2$ $+\ c\ E_4(3,4) \otimes E_4(3,4) \otimes Id_2$

This describes the dependency labeled (s_2).

Mixed Dependencies		
Description	Contribution	Related term
$I=\{1,2,3\}$, $K=\emptyset$, $I_0=\{3\}$,	for X_1:$Q_{X_1}(0)= E_4(4,4)$	$-d\ E_4(4,4) \otimes E_4(4,4) \otimes E_2(2,2)$
	$Q_{X_1}(1) = E_4(4,1)$	
$T=[4]\times[4]\times[2]$	$Q_{X_1}(2) = E_4(4,1)$	$+ d\ p\ E_4(4,1) \otimes E_4(4,1) \otimes E_2(2,1)$
Nb of destinations $\mid f \mid = 2$	for X_2:$Q_{X_2}(0)= E_4(4,4)$	$+ d\ (1-p)\ E_4(4,1) \otimes E_4(4,2) \otimes E_2(2,1)$
	$Q_{X_2}(1) = E_4(4,1)$	
	$Q_{X_2}(2) = E_4(4,2)$	$+ Id_4 \otimes Id_4 \otimes M_6$
	$Q_{X_3}(0) = -d\ E_2(2,2)$	
	$Q_{X_3}(1) = d\ p\ E_2(2,1)$	
	$Q_{X_3}(2) = d\,(1-p)\ E_2(2,1)$	
	$Q_{X_3} = M_6$	

$$M_1 = \begin{bmatrix} -a_1 & a_1 & 0 & 0 \\ 0 & -b_1 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad M_2 = \begin{bmatrix} -a_1 & a_1 & 0 & 0 \\ 0 & -b_1 & 0 & b_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad M_3 = \begin{bmatrix} -a_2 & a_2 & 0 & 0 \\ 0 & -b_2 & b_2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$M_4 = \begin{bmatrix} -a_2 & a_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad M_5 = \begin{bmatrix} -a_4 & a_4 \\ 0 & 0 \end{bmatrix} \quad M_6 = \begin{bmatrix} 0 & 0 \\ b_4 & -b_4 \end{bmatrix}$$

So, the generator is given by the following formula, after some modifications to eliminate the generalized operators:

$$Q_Z = M_1 \otimes Id_4 \otimes l_1(Id_2) + M_2 \otimes Id_4 \otimes l_2(Id_2)$$

$$+ Id_4 \otimes M_3 \otimes l_1(Id_2) + Id_4 \otimes M_4 \otimes l_2(Id_2)$$
$$+ Id_4 \otimes Id_4 \otimes M_5$$
$$- c\ E_4(3,3) \otimes E_4(3,3) \otimes Id_2 + c\ E_4(3,4) \otimes E_4(3,4) \otimes Id_2$$
$$-d\ E_4(4,4) \otimes E_4(4,4) \otimes E_2(2,2) + d\ p\ E_4(4,1) \otimes E_4(4,1) \otimes E_2(2,1)$$
$$+ d\ (1-p)\ E_4(4,1) \otimes E_4(4,2) \otimes E_2(2,1) + Id_4 \otimes Id_4 \otimes M_6$$

This is a typical example of a non-regular problem, where no factorization is possible within the formula.

We consider now the queuing example of Figure 5. We denote Ri the infinite matrix whose effect is the elementary right shift, and Le the infinite matrix whose effect is the elementary left shift. Id is also the infinite identity matrix.

$$Ri = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \qquad Le = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

The analysis yields:

Probabilistic Dependencies		
Description	Contribution	Related term
$I=\{1\}$, $K=\{2\}$, $T=\mathbb{N}$	$\overline{Q}_{X_1 \mid X_2}=(\lambda\ Ri\ + \nu\ n\ \ Le\ )_{n \in \mathbb{N}}$	$\overline{Q}_{X_1 \mid X_2} \otimes Id$
$I=\{2\}$, $K=\{1\}$, $T=\mathbb{N}$	$\overline{Q}_{X_2 \mid X_1}=(\alpha\ Ri\ + \nu\ 1(\ n < p\ )\ Le\ )_{n \in \mathbb{N}}$	$Id \otimes \overline{Q}_{X_2 \mid X_1}$

So, the generator is

$$Q_Z = \overline{Q}_{X_1 \mid X_2} \otimes Id + Id \otimes \overline{Q}_{X_2 \mid X_1}$$

We denote Ri_N, Le_N and Id_N respectively, the right shift, left shift and the identity matrix of dimension N. The analysis of the queuing example of Figure 6, yields the table:

Concurrent-firing Dependencies		
Description	**Contribution**	**Related term**
$I=\{1,2\}$, $K=\emptyset$, $T=[k]$ for all $k \in]0,N[$ Nb of destinations $\mid f \mid=2$	for X_1: $Q_{X_1}(0)=(-\mu-\nu)\, E_N(k,k)$ $Q_{X_1}(1)=\nu\, E_N(k,k+1)$ $Q_{X_1}(2)=\mu\, E_N(k,k-1)$ for X_2: $Q_{X_2}(0)=E_N(N-k,N-k)$ $Q_{X_2}(1)=E_N(N-k,N-k-1)$ $Q_{X_2}(2)=E_N(N-k,N-k+1)$	$(-\mu-\nu)\, E_N(k,k) \otimes E_N(N-k,N-k)$ $+\ \nu\, E_N(k,k+1) \otimes E_N(N-k,N-k-1)$ $+\ \mu\, E_N(k,k+1) \otimes E_N(N-k,N-k-1)$
$I=\{1,2\}$, $K=\emptyset, T=[0]$ Nb of destinations $\mid f \mid=1$	for X_1: $Q_{X_1}(0)=-\nu\, E_N(0,0)$ $Q_{X_1}(1)=\nu\, E_N(0,1)$ for X_2: $Q_{X_2}(0)=E_N(N,N)$ $Q_{X_2}(1)=E_N(N,N-1)$	$-\nu\, E_N(0,0) \otimes E_N(N,N)$ $+\ \nu\, E_N(0,1) \otimes E_N(N,N-1)$
$I=\{1,2\}$, $K=\emptyset, T=[N]$ Nb of destinations $\mid f \mid=1$	for X_1: $Q_{X_1}(0)=-\mu\, E_N(N,N)$ $Q_{X_1}(1)=\mu\, E_N(N,N-1)$ for X_2: $Q_{X_2}(0)=E_N(0,0)$ $Q_{X_2}(1)=E_N(0,1)$	$-\mu\, E_N(N,N) \otimes E_N(0,0)$ $+\ \mu\, E_N(N,N-1) \otimes E_N(0,1)$

This analysis leads to a numerable number of terms, but those can be reduced to the following:

$$Q_Z = -\mu \; [(Le_N \; Ri_N) \otimes Id_N] - \nu \, [Id_N \otimes (Ri_N \; Le_N)] + \mu \, (Le_N \otimes Ri_N) + \nu \, (Ri_N \otimes Le_N)$$