TRUE-CONCURRENCY PROBABILISTIC MODELS: MARKOV NETS AND A LAW OF LARGE NUMBERS

SAMY ABBES AND ALBERT BENVENISTE

True-concurrency Probabilistic Models: Markov Nets and a Law of Large Numbers[*]

Samy Abbes [**]   and Albert Benveniste [***]

Systèmes communicants
Projet DistribCom

Abstract: We introduce the model of Markov nets, a probabilistic extension of safe Petri nets under the true-concurrency semantics. This model builds upon our previous work on probabilistic event structures. We use the notion of branching cell for event structures and show that the latter provides the adequate notion of local state, for nets. We prove a Law of Large Numbers (LLN) for Markov nets—this constitutes the main contribution of the paper. This LLN allows characterizing in a quantitative way the asymptotic behavior of Markov nets.

Key-words: True concurrency, probabilistic models, safe Petri nets, Markov nets, law of large numbers

(Résumé : tsvp)

[*] A preliminary version of this paper was presented at the conference FOSSACS'05
[**] Institute for Systems Research, A. V. Williams Building, University of Maryland, College Park, MD 20742, USA
[***] IRISA/INRIA, Campus de Beaulieu, 35042 Rennes Cedex, France. E-mail: benveniste@irisa.fr, http://www.irisa.fr/distribcom/benveniste

Modèles probabilistes pour des systèmes concurrents : Réseaux Markoviens et Loi des Grands Nombres****

Résumé : On introduit le modèle des Réseaux markoviens, une extension probabiliste des Réseaux de Petri saufs avec sémantique de la concurrence vraie. Ce travail s'appuie sur un travail précédent concernant les structures d'événements probabilistes. On utilise la notion de cellule de branchement et on montre qu'un repliement de celles-ci fournit la bonne notion d'état local. On montre une loi des grands nombres sous concurrence, ce qui est la contribution principale de cet article.

Mots clés : Concurrence vraie, modèles probabilistes, réseaux de Petri saufs, réseaux Markoviens, loi des grands nombres

Contents

1 Introduction

This paper studies the model of *Markov nets,* a probabilistic model of safe Petri nets under true concurrency semantics. This means that traces, not firing sequences, are given a probability. This study belongs to the recently developed area of true concurrency probabilistic models, addressing both event structures and nets. Confusion-free probabilistic event structures were studied in [15, 14]. *Distributed probabilistic event structures* and *Markov nets* were investigated in [1], following an approach initiated in [6]; the latter approaches address event structures with confusion. This paper studies the model of *Markov nets,* a probabilistic extension of safe Petri nets under the true-concurrent semantics. The main contribution of the paper, w.r.t. to the conference version [5], is the complete study of the Law of Large Numbers for Markov nets.

We rely on previous work [3] on probabilistic event structures—probabilistic event structures are prime event structures whose space of maximal configurations is equipped with a probability.

In the above reference we introduced the class of so-called *locally finite* event structures, which encompasses in particular confusion-free event structures. Locally finite event structures are event structures with kind of finite confusion. For locally finite event structures, so-called *branching cells* generalize the cells of confusion-free event structures. Branching cells are sub-event structures that recursively localize the sources of conflict in an event structure.

Equipping branching cells with local probabilities and composing them together according to a generalized Bayes chain rule, defines a unique probability measure on the space of maximal configurations of the considered event structure. Probabilities obtained by this way are called *distributed*, for the following reason [3]: the so constructed probabilistic event structures are such that *"concurrent processes are independent in the probabilistic sense"*.

The goal of the present paper is to apply these constructions to the case of an event structure obtained by unfolding a safe Petri net. For this case, branching cells that are isomorphic when seen as event structures labeled by transitions from the net, are considered equivalent. The resulting equivalence classes are finitely many and are called *dynamic clusters*. Let us mention that dynamic clusters, as their name indicates, differ from usual clusters [8], which are statically defined on nets. Equipping the dynamic clusters of a safe Petri net with local probabilities yields *Markov nets.* Unfolding Markov nets gives raise to probabilistic event structures where equivalent branching cells are given the same local probability.

A central tool in probabilistic models is the *Law of Large Numbers* (LLN). It is the basis for the use of probabilistic models in statistics. It provides a quantitative counterpart to the central notion of fairness, for infinite executions of systems. Recall that the classical LLN for a sequence $(X_k)_{k>0}$ of independent random variables with identical distribution μ states as follows: for every nonnegative, real-valued, function f defined on the state space of X_k, the empirical means $\frac{1}{n} \sum_{k=1}^{n} f(X_k)$ converge with probability 1 to the mean value

$\int f(x)\mu(dx)$. This LLN extends to (recurrent) Markov chains, with μ being the invariant measure of the chain.

For Markov nets, the statement of a LLN is by itself doubly challenging, since:

1. there is no global time index k in the true-concurrency framework, and

2. it is unclear what the state space should be.

Regarding the second point, the normal guess that the local states would simply be the places of the net, does not work in general—places are not the right notion of local state in the probabilistic context. We will indeed show that

1. the "time elapsed while scanning a configuration" is adequately measured by the number of branching cells traversed, and

2. dynamic clusters (i.e., equivalence classes of branching cells) provide the right notion of state.

A third challenge is that, since configurations exhibit concurrency, the "progress of time" while scanning a configuration is not clearly defined: the different processes composing the configuration may progress freely as long as they do not need to synchronize. Our LLN requires that empirical means converge whatever way these different processes progress, and that the resulting limit should be unique.

Not every Markov net can obey such a LLN: in a net composed of two non interacting subnets, the two subnets can progress freely and no single "life time" can bound this progress. For such nets lacking synchronization, our LLN cannot hold. We thus finely characterize how much lack of synchronization can be tolerated, while still having the LLN valid: this is expressed as an integrability condition on a certain random variable that measures the lack of synchronization of the system. This condition is trivially satisfied for Markov chains, since Markov chains do not present any concurrency and thus "maximally synchronize".

To keep the present paper self-contained, the main results of [3] concerning branching cells of event structures and distributed probabilities are recalled. The paper is organized as follows: In §2, we introduce on a few toy examples the kind of randomization for Petri nets we consider in this paper. The Law of Large Number for true-concurrency systems is also discussed in an informal manner. This first section provides an intuitive introduction to probabilistic true-concurrency models. It only requires basic knowledge of Petri net theory. It is intended to both 1) researchers used to the interleaving semantics for probabilistic systems and who wish to have a short introduction to its true-concurrency counterpart, and 2) researchers comfortable with the true-concurrency semantics but not with its probabilistic counterpart.

The remaining of the paper assumes knowledge of fundamentals of unfolding theory of safe Petri nets [11]. We collect in §3 some basic notions concerning event structures, and fix the notations. The non-standard notions of stopping prefixes and branching cells are recalled from [3]. In §4, we recall the basic notion concerning probabilistic event structures

as well as the construction of distributed probabilities. Markov nets are introduced in §5, and we state the LLN in § 6. The proof of the LLN is the topic of §7. In Appendix A, we recall the classical statements of the LLN that we use.

2   Illustrative Examples

This section presents some toy examples to illustrate the issues encountered when randomizing Mazurkiewicz traces for a Petri net, and the solutions we propose. We first recall basic concepts of Mazurkiewicz trace theory. We explain on simple examples why *dynamic clusters* of nets must be considered, and their role in the randomization of nets. We also informally describe the Law of Large Numbers in this context.

True Concurrency and Mazurkiewicz Traces.   According to the true-concurrency approach to safe Petri nets, we do not distinguish a firing sequence $(\ldots, t, t', \ldots)$ from the firing sequence obtained by exchanging places t and t', whenever t and t' are transitions of the net that share no common resource. The equivalence classes—after taking the reflexive transitive closure of the above relation—are called the Mazurkiewicz traces of the net, or simply its traces for short [10, 11]. Figure 1 depicts the example of a safe Petri net. In this example, the firing sequence (bce) is equivalent to the sequence (cbe) since transitions b and c do not share common resources.
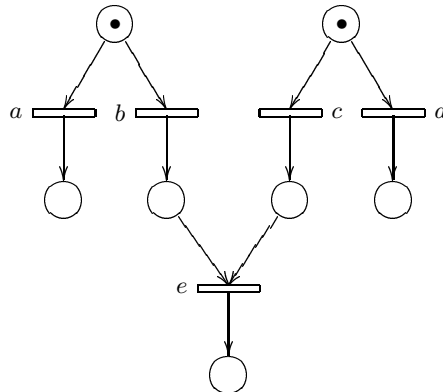
Figure 1: *Illustrating the true-concurrency semantics.*

Firing sequences of a safe Petri net are ordered by the prefix relation on words. This relation induces in turn an ordering on traces, so that trace σ precedes σ' if and only if there are firing sequences s and s', with s a prefix of s', representing σ and σ', respectively. The partial order on traces for the net of Figure 1 is depicted in Figure 2. On the other hand, each trace is seen itself as a partially ordered multi-set (*pomset*) of transitions, where two transitions are not comparable when they can be exchanged according to the above rule. In

the language of event structures, the elements that compose the pomset are called *events*. This is illustrated in Figure 2, (2), for the trace $(bce) = (cbe)$.

Figure 2: (1) *Ordering of the traces of the net depicted in Figure 1.* (2) The trace $(bce) = (cbe)$ as a partial order of labeled events.

Randomizing Maximal Traces. Traces are partially ordered sets of events. Therefore, classical approaches from the area of stochastic processes do not apply, since the latter assume a totally ordered time. We thus first discuss how Mazurkiewicz traces should be randomized.

We randomize a safe Petri net by defining a probability on its set of maximal traces. That is, we implicitly agree that *no blocking other than intrinsic deadlocks of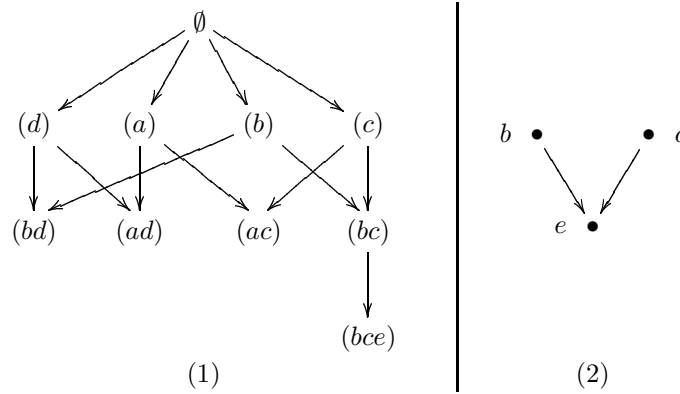 the net shall appear due to randomization*. Allowing the probability to weight non-maximal traces could be relevant for modeling reasons, for example, to express the possibility of failure due to a program crash. In the classical study of finite Markov chains, we indeed find this feature under the name of "killed processes" [13, p.25]. It is also known that killed trajectories can be made maximal by adding an additional dummy state in the state space, called the "cemetery". The same can be performed for nets, so we only focus on maximal traces throughout this paper. For the example depicted in Figure 1, maximal traces are the extremal bottom points of Figure 2 (1), i.e., (bd), (ad), (ac) and (bce). Making the net probabilistic amounts thus to defining a probability $\mathbb{P}$ such that:

$$\mathbb{P}(bd) + \mathbb{P}(ad) + \mathbb{P}(ac) + \mathbb{P}(bce) = 1.$$

Observe that, if we sum these probabilities, not on traces but on firings sequences, the result will exceed 1. This means that considering true-concurrency semantics has a significant impact on the construction of probabilities.

Concurrency and Probability. Our constant philosophy in the construction of proba-
bilistic nets is the following: as much as possible, parallel processes shall be made indepen-
dent in the probabilistic sense. When considering a Petri net as a distributed system with
a distributed state, this requirement is quite natural. In turn, synchronization is a major
source of difficulty. Indeed, when components do not not interact at all, we can simply ran-
domize each component separately, and make them probabilistically independent by decree.
In general, however, processes are "parallel" for a limited amount of "time", then synchro-
nize, which results in breaking the parallelism. We shall give a precise formulation of the
latter claim, by defining a decomposition of processes through locally parallel components.
With this decomposition of processes at hand, we shall proceed with their randomization,
by enforcing probabilistic independence of locally parallel components. This is not a trivial
task, but is successfully achieved by using classical tools from Measure theory. The resulting
theory is presented in detail in [3], for the model of event structures.

Dynamic Clusters. Consider first the two nets depicted in Figure 3, (a) and (b). Net
in (a) offers three possibilities: either firing transition a, or firing transition b, or firing no
transition at all. Since we only consider maximal traces, as explained above, we deny to
this net the right of doing nothing, so that it must eventually fire transition a or b. This is
a simple coin tossing, that occurs with a certain probability, say $\mu_1(a)$ versus $\mu_1(b)$.

Consider next the net of Figure 3, (b). Again, since we consider only maximal traces
for randomization, the net has exactly two possibilities: either firing transition d, or firing
concurrently transitions c and e. We have again a coin tossing. This time, however, the
tossing does not involve single transitions, but rather certain groups of transitions. We
shall thus weight the occurrence of (ce) with some probability $\mu_2(ce)$, while $\mu_2(d)$ is the
probability that (d) occurs instead of (ce). Remark that, although transitions c and e are
concurrent, they are not independent, since c occurs if and only if e occurs.

Let us now investigate the effect of synchronization. The net of Figure 3 (c) collects the
two previous nets and adds some new elements whose execution depends on the previous
execution of the first two nets. In this discussion, we shall refer to the first two nets,
considered separately, as *local nets*, and to their executions as *local executions*. In contrast,
the net of Figure 3 (c) is referred to as the *global net*. Because the two local nets do not
interact at all, the local execution of one of them does not disturb the execution of the other
one. Hence the local executions are *concurrent* and do not interact.

Suppose that each local net is driven by some random local agent, and that these local
agents do not communicate with each other. It is then natural, from the probabilistic point
of view, to assume that the actions of local agents are independent in the probabilistic sense.
Observe that a maximal trace of the global net induces, by restriction, a maximal trace of
each local net. This is important, since we have only specified how to randomize the maximal
traces of the two local nets.

Consequently, we know at this point how to randomize the "beginning" of the executions
of the global net. If μ_1 and μ_2 denote respectively the probabilities attached to each local
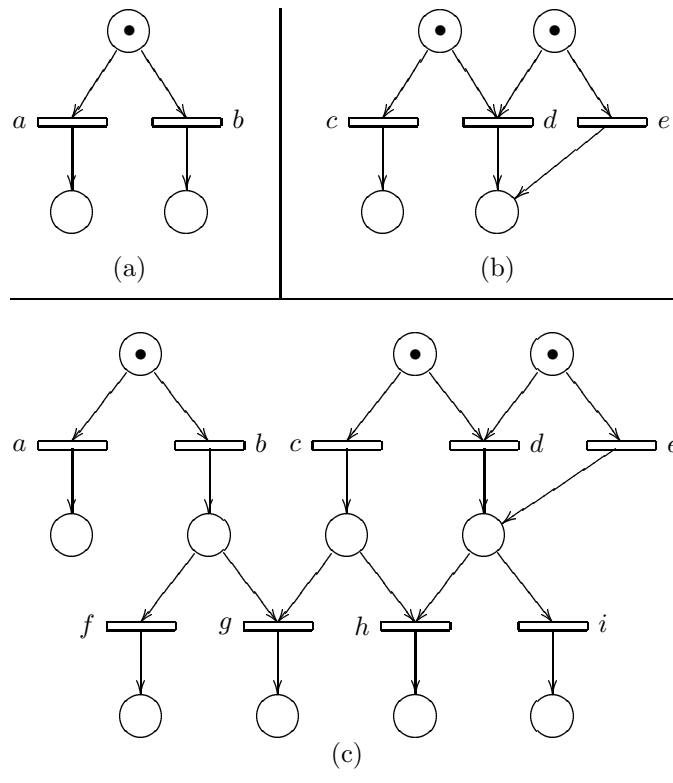
Figure 3: *Illustrating local parallelism and synchronization of safe Petri nets.*

net, so that

$$\mu_1(a) + \mu_1(b) = 1, \quad \mu_2(ce) + \mu_2(d) = 1,$$

and if $\mathbb{P}$ denotes the global probability constructed so far, we have for example:

$$\mathbb{P}(a\,ce) = \mu_1(a) \times \mu_2(ce), \quad \mathbb{P}(b\,ce) = \mu_1(b) \times \mu_2(ce).$$

In the above equation, the left members actually concern maximal executions of the global net: they must be understood as, for example: "$\mathbb{P}(a\,ce)$ is the probability that a maximal execution contains (ace)".

What happens next? Consider for example the case of (ace) having fired. The resulting net is depicted in Figure 4 (a). The transitions that are not enabled are depicted with dashes, and we omit the transitions that have already fired. The only enabled transitions are transitions h and i. We shall thus consider some probability μ_3, that describes the weight of h against i, *in the context of h and i competing alone.* By the chain rule, the complete probabilities of $(aceh)$ and of $(acei)$ can now be computed by:

$$\mathbb{P}(a\,ce\,h) = \mu_1(a) \times \mu_2(ce) \times \mu_3(h), \quad \mathbb{P}(a\,ce\,i) = \mu_1(a) \times \mu_2(ce) \times \mu_3(i).$$
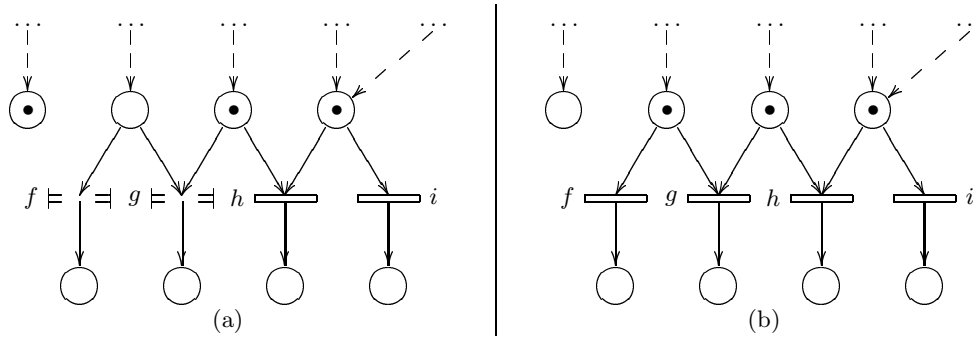
Figure 4: *Nets resulting from different partial executions of the net depicted in* Figure 3 (c).

Instead of (a) and (ce) being the executions of the two first local nets, assume that, for example, (b) and (ce) have fired. The resulting net is depicted in Figure 4 (b). Now all transitions f, g, h and i are enabled. Hence h and i still compete for firing, but they do not compete in the same context as previously observed, since f and g also take part to the competition. We shall thus consider yet another local probability μ_4, that randomizes the maximal traces of the net depicted in Figure 4 (b), i.e., such that:

$$\mu_4(fh) + \mu_4(fi) + \mu_4(gi) = 1.$$

We again use the chain rule to compute the probability of, say, $(bcefh)$:

$$\mathbb{P}(b\,ce\,fh) = \mu_1(b) \times \mu_2(ce) \times \mu_4(fh).$$

Finally, in case of (*ad*) or (*bd*) firing in the first two local nets, we would consider also the additional local nets consisting of transition *i* only, and of transition *f* only. Both nets are necessarily equipped with trivial probabilities, i.e., with probabilities giving weight 1 to the unique possible transition.

With the recursive decomposition of traces described so far, we have reached all maximal configurations of the global net and we know how to compute the probability of each maximal configuration. We leave to the reader as an exercise to check on this example that the probability defined by this way sums indeed to 1 on the set of all maximal configurations of the global net, by using the fact that μ_1 and μ_2 both sum to 1.

The different local nets encountered in the course of all possible executions of the net are called *dynamic clusters*. The global net we have studied has 6 dynamic clusters: the two nets of Figure 3 (a) and (b), the subnet obtained from Figure 4 (a) by keeping only transitions *h* and *i*, the one obtained from Figure 4 (b) by keeping only transitions *f*, *g*, *h* and *i*, and finally the two trivial nets with single transitions *f* and *i*.

The decomposition through these subnets is indeed *dynamic*, since a same transition may occur in different dynamic clusters, according to the context. This was the case, for example, for the transitions *h* and *i*. This fact conforms with the intuition that concurrent systems shall not be statically decomposed, but may split and join in different manners, according to the actual execution. We summarize what we have obtained so far:

1. We randomize maximal Mazurkiewicz traces of safe Petri nets.

2. Maximal traces are decomposed as the juxtaposition of maximal traces of dynamic clusters.

3. To each dynamic cluster, we attach an agent that randomizes the maximal traces of this cluster.

4. Concurrent dynamic clusters do not interfere with each other. They can thus be made independent in the probabilistic sense.

The probabilistic systems constructed in this way are called *Markov nets.*

Regarding the Law of Large Numbers. Recall the classical LLN for a sequence $(X_k)_{k>0}$ of independent random variables with identical distribution μ: for every non-negative, real-valued, function f defined on the state space of X_k, the empirical means $\frac{1}{n} \sum_{k=1}^{n} f(X_k)$ converge with probability 1 to the mean value $\int f(x)\mu(dx)$. This LLN extends to (recurrent) Markov chains, with μ being the invariant measure of the chain. We shall now discuss the LLN for Markov nets. Of course, such a LLN is relevant only for nets having infinite configurations.

An example of such a net is shown in Figure 5 and its dynamic clusters are depicted in Figure 6. Note that these clusters overlap, revealing their dynamic nature. For example, transition *d* belongs to both clusters $\mathbf{s}_4$ and $\mathbf{s}_5$; in fact, when a configuration *v* traverses *d*, then it traverses either cluster $\mathbf{s}_4$ or cluster $\mathbf{s}_5$ but it cannot traverse both.

As observed in the introduction, the first difficulty consists in finding the adequate notion of state, for Markov nets. Since dynamic clusters capture the choices made when constructing a maximal trace, they are a natural candidate for the notion of (local) state. Therefore, our "state functions" will be real-valued functions defined on the set Σ of dynamical clusters. The next issue is that of finding the counterpart of the time index k in empirical means $\frac{1}{n}\sum_{k=1}^{n} f(X_k)$. We regard k as indexing the successive outcomes of the random choices made while drawing the considered trajectory. Therefore, a natural counterpart of k is the number of dynamic clusters traversed while constructing the considered configuration.
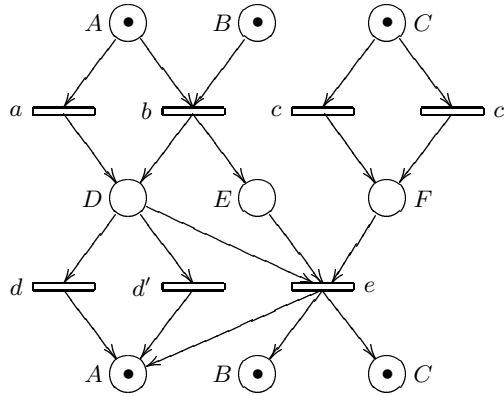
Figure 5: *A safe net with infinite executions. Places having the same name are to be superimposed.*

The above analysis suggests the following form for the empirical means in our case:

$$M(f, v) = \frac{\sum_{\mathbf{s} \in \Sigma: \, v \text{ traverses } \mathbf{s}} f(\mathbf{s})}{\text{number of dynamic clusters } \mathbf{s} \text{ traversed by } v} \, , \tag{1}$$

where f is a state function and v is some finite trace. It must be noted that a same cluster is traversed several times by the finite trace v, such traversals are therefore counted with their multiplicities in the above summation $\sum_{\mathbf{s}}$.

Consider in particular (1) with $f(\mathbf{s}) = \mathbf{1}_{\mathbf{s}_*}(\mathbf{s})$, where $\mathbf{1}_{\mathbf{s}_*}(\mathbf{s}) = 1$ if $\mathbf{s} = \mathbf{s}_*$, and $= 0$ otherwise. Then,

$$M(\mathbf{1}_{\mathbf{s}_*}, v) = \frac{\text{number of occurrences of } \mathbf{s}_* \text{ seen by } v}{\text{number of dynamic clusters } \mathbf{s} \text{ traversed by } v}.$$

Assume for a while that, for each maximal trace ω and each dynamic cluster $\mathbf{s}_*$,

$$\alpha(\mathbf{s}_*) =_{\text{def}} \lim_{v \nearrow \omega} M(\mathbf{1}_{\mathbf{s}_*}, v) \tag{2}$$

exists and does not depend on ω. Since, for every v, $\sum_{\mathbf{s}} M(\mathbf{1}_{\mathbf{s}}, v) = 1$, the $\alpha_{\mathbf{s}}$ sum up to 1 and therefore define a probability on the finite set Σ, we call it the *stationary measure* of
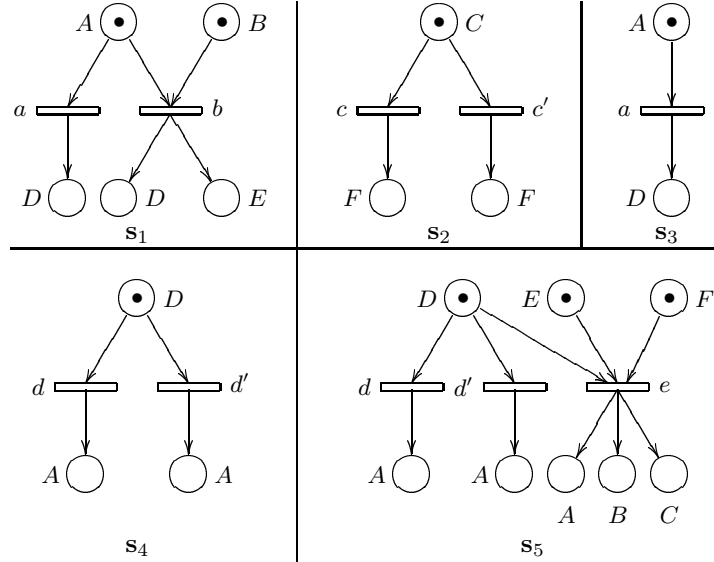
Figure 6: *Dynamic clusters* $s_1, \ldots, \mathbf{s}_5$, *of the net depicted in Figure 5. The figure depicts the unfolded form of the clusters. Here letters indicate the labels, not the names of the node, and places with same labels are not to be superimposed.*

the Markov net. Then, by linearity, for every nonnegative state function f,

$$\lim_{v \nearrow \omega} M(f, v) = \sum_{\mathbf{s} \in \Sigma} f(\mathbf{s}) \alpha(\mathbf{s})$$

exists and does not depend on ω. This provides us with the desired LLN. Thus it is enough proving the LLN for the special case (2).

A direct application of this LLN yields in particular the asymptotic ratio:

$$\lim_{v \nearrow \omega} \frac{\text{number of occurrences of } \mathbf{s} \text{ in } v}{\text{number of occurrences of } \mathbf{s}' \text{ in } v} = \frac{\alpha(\mathbf{s})}{\alpha(\mathbf{s}')},$$

for $\mathbf{s}, \mathbf{s}'$ two dynamic clusters (see Figure 6 for an illustration of the clusters). Now, a more natural question would be to evaluate the asymptotic ratio:

$$\lim_{v \nearrow \omega} \frac{\text{number of occurrences of } t \text{ in } v}{\text{number of occurrences of } t' \text{ in } v}, \tag{3}$$

for t, t' two transitions of the net. Unfortunately, transitions are not in bijection with dynamic clusters (as shown by Figure 6), and therefore the above asymptotic ratio is not within the scope of our above LLN.

Suppose now we were able to prove a LLN for "extended" state functions of the form $f(\mathbf{s}, \omega_{\mathbf{s}})$, where $\mathbf{s}$ is a dynamic cluster as before, and $\omega_{\mathbf{s}}$ is a maximal trace of $\mathbf{s}$, seen as an event structure. Corresponding empirical means would be:

$$M(f, v) = \frac{\sum_{\mathbf{s} \,\in\, \Sigma \,:\, v \text{ traverses } \mathbf{s}} f(\mathbf{s}, \omega_{\mathbf{s}})}{\text{number of dynamic clusters } \mathbf{s} \text{ traversed by } v},$$

where $\omega_{\mathbf{s}}$ is the maximal trace of $\mathbf{s}$ seen by v when traversing $\mathbf{s}$, and the traversals of $\mathbf{s}$ by v are counted with their multiplicities, as in (1). Then,

$$\mathbf{1}_t(\mathbf{s}, \omega_{\mathbf{s}}) =_{\text{def}} \begin{cases} 1, & \text{if } \omega_{\mathbf{s}} \text{ visits } t \\ 0, & \text{otherwise} \end{cases}$$

would yield an extended state function such that

$$\frac{M(\mathbf{1}_t, v)}{M(\mathbf{1}_{t'}, v)} = \frac{\text{number of occurrences of } t \text{ in } v}{\text{number of occurrences of } t' \text{ in } v},$$

so that our extended LLN would encompass asymptotic ratios of the form (3). We shall indeed prove such an extended LLN and therefore solve the problem of evaluating asymptotic ratios of the form (3).

So far we have used the expression

$$\lim_{v \nearrow \omega} M(f, v) \tag{4}$$

without care. Indeed, trace v can grow to ω in many different ways. If, for example, we erase, in the net of Figure 5, the places B and E, and the arcs $D \to e$ and $e \to A$, then the two remaining noninteracting subnets can progress freely. And it is unclear whether the limit (4) is well defined, since it could very well depend on the way the two components of v grow. Not surprisingly, the LLN we shall establish assumes that the different local processes of the net do synchronize "frequently enough". In turn, we will show that, with this assumption, the limit (4) is well defined and does not depend on the way v can grow to ω.

3   Locally Finite Event Structures

In this section we recall basic definitions concerning event structures and we fix the notations. We introduce the three main notions of *locally finite* event structure, *stopping prefix,* and *future* of a configuration. The latter notion has already been considered under different names by several authors. However, we use it in a novel, systematic way for the analysis of event structures.

3.1   Event Structures and Net Unfoldings.

Throughout this paper, the term "event structure" means *prime* event structure.

An event structure [11] is a triple $(E, \preceq, \#)$ satisfying the following properties. $(E, \preceq)$ is a partial order, at most countable, whose elements are called events. $\#$ is the *conflict* relation on E; it is a binary, symmetric and irreflexive relation, that satisfies the following axiom: $\forall x, y, z \in E$, $x \# y$ and $y \preceq z$ together imply $x \# z$. We also assume that, for every $e \in E$, the downward closure of e, defined by $\lceil e \rceil = \{f \in E : f \preceq e\}$ is a finite subset of E. We identify, with a slight abuse of notations, the set E and the event structure $(E, \preceq, \#)$. Finally, the *concurrency relation* is the binary relation on E, denoted by **co**, and defined by $\mathbf{co} = (E \times E) \setminus (\# \cup \preceq \cup \succeq)$. We say that E is a *tree of events*, or shortly, a *tree*, if $\mathbf{co} = \emptyset$.

A subset A of E is said to be a *prefix* if it is downward closed, i.e., if $\lceil e \rceil \subseteq A$ for all $e \in A$. A prefix v is called a *configuration* of E if it is *conflict-free*, i.e., if $\# \cap (v \times v) = \emptyset$. Configurations are partially ordered by inclusion, and we denote by $\mathcal{V}_E$ the poset of *finite* configurations of E. Two configurations are said to be *compatible* if their set-theoretic union is conflict-free. We denote by Ω_E the set of maximal configurations of E—this set is nonempty, since chain of configurations has an upper bound.

A subset $F \subseteq E$ defines a sub-event structure $(F, \preceq_F, \#_F)$ of E with causality and conflict relations inherited by:

$$\preceq_F = \preceq \cap (F \times F), \quad \#_F = \# \cap (F \times F),$$

and we shall freely write F, $\mathcal{V}_F$, and Ω_F to denote this event structure and its set of finite and maximal configurations, respectively.

All the material introduced in this paper regarding event structures is intended to apply to the case of an event structure obtained by unfolding a safe Petri net. See [11] for a basic reference on this topic. That is, a safe Petri net $\mathcal{N} = (P, T, F, m_0)$ is given, with set of places P, set of transitions T, flow relation F and initial marking m_0. The unfolding of $\mathcal{N}$ is an event structure E equipped with a labeling mapping $\lambda : E \to T$. The configurations of E determine a family of pomsets (partially ordered multi-sets), via the labelling λ. The unfolding (E, λ) is characterized by the fact that traces of $\mathcal{N}$ correspond exactly to pomsets induced by the configurations of E. We will not use the labeling mapping until §5, and consider instead E as an abstract event structure. However, since E arises from the unfolding of a safe net, it is legitimate to assume that E is such that any finite configuration enables only finitely many events. Formally:

Assumption 1   *For every $v \in \mathcal{V}_E$, $\mathrm{Min}\big(E^v\big)$ contains finitely many events.*

3.2   Future of a Configuration

For v a finite or infinite configuration of E, we consider the following subset of E:

$$E^v =_{\mathrm{def}} \{e \in E \setminus v : \forall e' \in v, \neg(e \# e')\}$$

We call the associated sub-event structure the *future* of v. It is clear that the map:

$$w \in \mathcal{V}_{E^v} \mapsto v \cup w \tag{5}$$

is one-to-one and onto, from the set of configurations of E^v to the set of configurations of E that contains v. The map (5) describes the *concatenation* of v with configurations of E^v. This map is also order-preserving, so that it maps maximal elements to maximal elements; that is, the set of maximal configurations of E^v is mapped onto the set of maximal configurations of E that contain v.

In order to distinguish $v \cup w$, for $v \in \mathcal{V}_E$ and $w \in \mathcal{V}_{E^v}$ from the usual set-theoretic union of compatible configurations of E, we introduce the following special notations for the concatenation and for its inverse:

$$\begin{aligned} v \oplus w =_{\text{def}} v \cup w, \quad &\text{defined for } v \in \mathcal{V}_E \text{ and } w \in \mathcal{V}_{E^v}, \\ u \ominus v =_{\text{def}} u \setminus v, \quad &\text{defined for } v \in \mathcal{V}_E \text{ and } u \in \mathcal{V}_E \text{ such that } v \subseteq u, \end{aligned} \tag{6}$$

so that we have, whenever these are well-defined:

$$v \oplus w \in \mathcal{V}_E, \qquad u \ominus v \in \mathcal{V}_{E^v} \,.$$

In the context of net unfoldings, the operation $\oplus$ defined above corresponds to the concatenation of traces of the net.

3.3   Stopping Prefixes

A central concept in defining probabilities is the notion of choice. Choice is therefore a key concept in this paper. It is captured by the notion of minimal conflict we recall next. The *minimal conflict* relation $\#_\mu$ on E is defined by:

$$\forall e, e' \in E, \quad e \,\#_\mu\, e' \iff (\lceil e \rceil \times \lceil e' \rceil) \cap \# = \{(e, e')\}.$$

Definition 3.1 (stopping prefix)   *A subset $B \subseteq E$ is called a* stopping prefix *of E if:*

1. *B is a prefix of E;*

2. *B is closed under minimal conflict.*

Stopping prefixes form a complete lattice, with $\emptyset$ and E as minimal and maximal elements. Probabilistic constructions consist in randomizing choices and should be therefore based on stopping prefixes. Hence the following notion is natural in this context:

Definition 3.2 (locally finite event structure)   *E is called* locally finite *if for each event e of E, there exists a* finite *stopping prefix containing e.*

Locally finite event structures have not been considered by authors so far. The following condition is implicitly assumed throughout this paper:

Assumption 2 *E is locally finite.*

It is easily checked that, if v is any configuration of E, e an event of E^v, and B a finite stopping prefix containing e, then $B \cap E^v$ is a finite stopping prefix of E^v containing e. As a consequence, every future E^v is locally finite. Stopping prefixes satisfy the following property (see [1, Ch.3, I-3.*1*]):

Lemma 3.1 *If B is a stopping prefix of E, then:*

$$\Omega_B = \{\omega \cap B \mid \omega \in \Omega_E\}. \tag{7}$$

This property, which is not satisfied by arbitrary prefixes (draw an example!), motivates introducing the following definition:

Definition 3.3 (stopped configurations) *A configuration v is called a* stopped configuration *of E if there is a stopping prefix B such that $v \in \Omega_B$.*

REMARK. Use Lemma 3.1 and the fact that E is locally finite to show that a stopped configuration v if finite if and only if there is a *finite* stopping prefix B such that $v \in \Omega_B$.

3.4 Branching Cells and R-stopped Configurations

We seek for the following two-steps procedure for constructing probabilistic event structures: 1/ consider stopped configurations as "elementary process" for randomization, and 2/ use the concatenation of stopped configurations in order to randomize processes in an incremental way.

Unfortunately, the class of stopped configurations is not closed under concatenation. That is, if v is a stopped configuration of E, and w a stopped configuration of the future E^v, then $v \oplus w$ is not stopped in E in general[1] (see an example in §3.5). This is why we consider the closure of stopped configurations under concatenation. The configurations reached by this way are called *recursively stopped* (R-stopped for short).

Definition 3.4 (R-stopped configurations) *A configuration v is said to be R-stopped in E if there exists a (finite or infinite) nondecreasing sequence $(v_n)_{0 \leq n < N}$ of configurations, $N \leq +\infty$, satisfying the following conditions:*

1. $v_0 = \emptyset$, $v = \bigcup_{0 \leq n < N} v_n$, and

2. $v_n \ominus v_{n-1}$ is a finite stopped configuration of the future $E^{v_{n-1}}$, for every $0 < n < N$.

The set of all finite R-stopped configurations is denoted by $\mathcal{W}_E$, or simply $\mathcal{W}$ if no confusion can occur.

---

[1] This is particular to systems with concurrency, and more precisely to systems with confusion: for confusion-free event structures, stopped configurations are closed under concatenation.

R-stopped configurations are thus characterized by the existence of a recursive decomposition through stopped configurations. Such decomposition is by no way unique. We shall however focus on *minimal* decompositions. Branching cells are introduced to this end.

Definition 3.5 (initial stopping prefix, branching cell)   *A stopping prefix B is called* initial *in E if $\emptyset$ is the only stopping prefix of E strictly contained in B. Call* branching cell *of E any initial stopping prefix of some E^v, for v ranging over $\mathcal{W}_E$. The set of all branching cells of E is denoted by X_E (or simply X) and branching cells are generically denoted by the symbol x. For $v \in \mathcal{W}_E$, the set of branching cells that are initial stopping prefixes of E^v is denoted by*

$$\delta_E(v),$$

or simply by $\delta(v)$ when no confusion can occur.

$\delta_E(v)$ must be understood as the set of branching cells "enabled" by v. The following theorem collects the different key properties of branching cells (proofs are found in [1, Ch.3], or in [3]).

Theorem 3.1

1. **Existence and finiteness of branching cells.** *Every branching cell of E is a finite sub-event structure of E. For every $v \in \mathcal{W}_E$, $\delta_E(v)$ is empty if and only if v is maximal in E.*

2. **Concurrent branching cells.** *Let v be any finite R-stopped configuration. Different branching cells in $\delta_E(v)$ are disjoint and concurrent, the latter meaning that:*

$$\forall x, y \in \delta_E(v), \quad x \neq y \implies \forall (e, f) \in x \times y, \quad e \textbf{ co } f.$$

   This has the two following consequences:

   (a) *For $v \in \mathcal{W}_E$, let B be a stopping prefix of E^v of the form $B = \bigcup_{x \in \xi} x$, where ξ is any subset of $\delta_E(v)$. Then configurations and maximal configurations of B respectively decompose as the following products:*

   $$\mathcal{V}_B = \prod_{x \in \xi} \mathcal{V}_x, \qquad \Omega_B = \prod_{x \in \xi} \Omega_x. \tag{8}$$

   (b) *For any $v \in \mathcal{W}_E$, $\delta_E(v)$ is finite (i.e., there are finitely many concurrent branching cells), and even bounded by some constant K.*

3. **Covering map.** *For v any R-stopped configuration, there exists a non-decreasing sequence of configurations $(v_n)_{0 \leq n < N}$, with $N \leq +\infty$, and a sequence of branching cells $(x_n)_{0 < n < N}$ with $x_n \in \delta_E(v_n)$ for all n, such that:*

   (a) *$v_0 = \emptyset$, $v = \bigcup_{0 \leq n < N} v_n$, and*

(b) $v_n \ominus v_{n-1}$ is a maximal configuration of x_n for all $0 < n < N$.

Branching cells $\{x_n, 0 < n < N\}$ are pairwise disjoint. If $(v'_n)_{0 \leq n < N'}$ is another such decomposition, with associated branching cells $(x'_n)_{0 < n < N'}$, then we have the equality of sets:

$$\{x_n, 0 < n < N\} = \{x'_n, 0 < n < N'\}.$$

In particular, $N = N'$. We call covering map of E *the map Δ_E defined by:*

$$\Delta_E(v) = \{x_n, 0 < n < N\}, \tag{9}$$

which only depends on v and on E.

4. ***Covering maps in stopping prefixes.*** *For any stopping prefix B of E, and for any configuration u of B, u is R-stopped in B if and only if u is R-stopped in E. In this case the covering maps Δ_E and Δ_B satisfy:*

$$\Delta_B(v) = \Delta_E(v).$$

In particular, $X_B \subseteq X_E$, with X denoting the sets of branching cells.

5. ***Concatenation and subtraction of R-stopped configurations; covering the future.*** *The class of R-stopped configurations is stable under concatenation and under subtraction:*

$$\forall u \in \mathcal{W}_E, \quad \forall v \in \mathcal{W}_{E^u}, \quad u \oplus v \in \mathcal{W}_E \,,$$
$$\forall u \in \mathcal{W}_E, \quad \forall w \in \mathcal{W}_E, \quad u \subseteq w \Rightarrow w \ominus u \in \mathcal{W}_{E^u} \,.$$

For $u \in \mathcal{W}_E$ and $v \in \mathcal{W}_{E^u}$, the covering maps Δ_E and Δ_{E^u} in the future E^u satisfy:

$$\Delta_E(u \oplus v) = \Delta_E(u) \cup \Delta_{E^u}(v), \quad \Delta_E(u) \cap \Delta_{E^u}(v) = \emptyset. \tag{10}$$

In particular, $X_{E^u} \subseteq X_E$ for any $u \in \mathcal{W}_E$. Finally, compatible R-stopped configurations form a lattice.

It must be noted that, except for the property that branching cells are finite, *all results stated above remain valid without the local finiteness assumption.* Further results requiring local finiteness will be stated in §3.6. First, we detail some examples.

3.5   Examples of Decompositions

For all examples of this paper, we write (abc) to denote the configuration $\{a, b, c\}$. To depict event structures, we use arrows for representing the causality and zigzag arcs for the minimal conflicts, as in Figure 7.

Figure 7 depicts with dashed frames the nonempty stopping prefixes of an event structure E. x and y are the two initial stopping prefixes of E. In this example, configuration
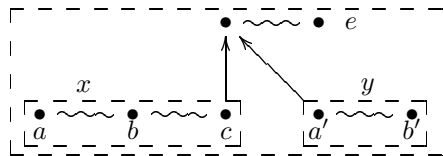
Figure 7: **Illustrating stopping prefixes and initial stopping prefixes.** *Causality is depicted by arrows, minimal conflicts are depicted with zigzag arcs. All nonempty stopping prefixes are shown with dashed frames. x and y are the two initial stopping prefixes of the event structure. Observe that e is a minimal event although it does not belong to any initial stopping prefix.*

(a) is not R-stopped. Indeed, otherwise (a) would be maximal in x, which is not $((ac)$ is maximal in $x)$. Hence, *there exists in general finite configurations that are not R-stopped. A fortiori*, such configurations are not stopped.

Still in the event structure of Figure 7, let v be the configuration $v = (aca')$. v is maximal in $B_0 = x \cup y$, hence v is stopped and in particular v is R-stopped. We show that the covering $\Delta_E(v)$ is given by $\Delta_E(v) = \{x, y\}$. Since v is stopped in B_0, it follows from Lemma 3.1 that $v_x = v \cap x$ is stopped in x. Hence, by point 5, we have:

$$\Delta_E(v) = \Delta_E(v_x) \cup \Delta_{E^{v_x}}(v \ominus v_x).$$

By point 4 of Theorem 3.1 $\Delta_E(v_x) = \Delta_x(v_x)$. Since x, as an event structure, is the only nonempty stopping prefix of itself, it is clear that $\Delta_x(v_x) = \{x\}$. Hence we have $\Delta_E(v) \supseteq \{x\}$. Symmetrically, we also obtain that $\Delta_E(v) \supseteq \{y\}$. But $v = (v \cap x) \cup (v \cap y)$, so we are done: $\Delta_E(v) = \{x, y\}$.

More generally, we retain that *if v has the form $v = \bigcup_{x \in \xi} v_x$, where ξ is a set of initial stopping prefixes, and v_x is maximal in x for every $x \in \xi$, then the covering $\Delta_E(v)$ is given by $\Delta_E(v) = \xi$.* Although quite intuitive, this result is not obvious from the only definition of the covering map.

As another example, consider the event structure E depicted in Figure 8. We shall determine the coverings of the two maximal configurations $\omega_1 = (ad)$ and $\omega_2 = (bce)$ of E.

Let $x = \{a, b\}$ be the unique initial stopping prefix of E. Then $\Delta_E(\omega_1)$ and $\Delta_E(\omega_2)$ both contain x. Figure 8 (2) and (3) respectively depict the futures $E^{(b)}$ and $E^{(a)}$ of configurations (b) and (a), with the associated initial stopping prefixes:

$$\delta_E(b) = \{z, z'\}, \quad \delta_E(a) = \{y\},$$

with $z = \{c\}$, $z' = \{e\}$, $y = \{c, d, e\}$. Hence we obtain $\Delta(\omega_1) = \{x, y\}$ and $\Delta(\omega_2) = \{x, z, z'\}$. This example shows that *branching cells of an event structure may overlap*, although branching cells in a same covering $\Delta_E(v)$ shall not overlap as stated by point 3 of Theorem 3.1.
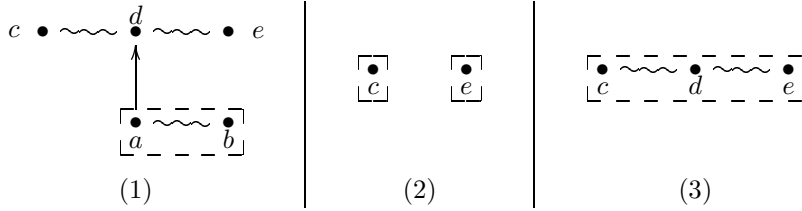
Figure 8: ***Illustrating the decomposition of R-stopped configurations.*** *(1), an event structure with two maximal configuration $\omega_1 = (ad)$ and $\omega_2 = (bce)$. (2), the future of (b), and (3), the future of (a). Initial stopping prefixes are depicted with dashed frames.*

3.6   Normal Decomposition of Maximal Configurations

We shall now give a systematic way to decompose maximal configurations, we call it the *normal decomposition.* This result deeply depends on the local finiteness assumption. This normal decomposition is of interest per se. It will be instrumental in proving the Law of Large Numbers.

For E a (locally finite) event structure, we define the *max-initial stopping prefix* of E as the upper bound of initial stopping prefixes. That is, denoting the max-initial stopping prefix of E by $B_0(E)$:

$$B_0(E) =_{\mathrm{def}} \bigcup_{x \in \delta_E(\emptyset)} x\,,$$

where x ranges over the set of initial stopping prefixes of E. Observe that, since initial stopping prefixes are finitely many according to Theorem 3.1, point 2b, and since each branching cell is finite by according to Theorem 3.1, point 1, the max-initial stopping prefix $B_0(E)$ is itself a finite stopping prefix.

We define, for ω a maximal configuration of E, the *normal decomposition* of ω as the following pair of sequences, $\big(V_n(\omega)\big)_{n \geq 0}$ and $\big(Z_n(\omega)\big)_{n \geq 1}$:

$$V_0 = \emptyset, \qquad n > 0, \quad \begin{cases} Z_n(\omega) = \omega \cap B_0\big(E^{V_{n-1}(\omega)}\big) \\ V_n(\omega) = V_{n-1}(\omega) \oplus Z_n(\omega) \end{cases} \tag{11}$$

Theorem 3.2 ([3])   *For every $\omega \in \Omega$, the sequence $\big(V_n(\omega)\big)_{n \geq 0}$ is a nondecreasing sequence of finite R-stopped configurations satisfying: $\sup_n V_n(\omega) = \omega$.*

See an interpretation of Theorem 3.2 by means of σ-algebras and probability in §7.1, Lemma 7.1.

4 Application of Branching Cells to Probabilistic Event Structures

4.1 Background on Probability Spaces

We first recall some basic notions from Probability theory, see for example [7].

Measurable Spaces and Measurable Mappings. Let U be a nonempty set. A σ-*algebra* on U is a collection $\mathfrak{F}$ of subsets of U such that $\emptyset \in \mathfrak{F}$, and $\mathfrak{F}$ is closed under complement and under countable union. The pair $(U, \mathfrak{F})$ is said to be a measurable space. The sets $A \in \mathfrak{F}$ are called $\mathfrak{F}$-measurable, or simply the measurable subsets of U. If $(U, \mathfrak{F})$ and $(V, \mathfrak{G})$ are two measurable spaces, a mapping $\varphi : U \to V$ is said to be a *measurable mapping* if $\varphi^{-1}(A) \in \mathfrak{F}$ for any $A \in \mathfrak{G}$. We usually adopt the folklore of probability, where measurable mappings are called *random variables*. A bijective mapping $\varphi : U \to V$ is said to be *bi-measurable* if both φ and φ^{-1} are measurable.

Let U be a nonempty set. For any collection $\mathcal{F}$ of subsets of U, there is a smallest σ-algebra $\mathfrak{F}$ that contains $\mathcal{F}$. $\mathfrak{F}$ is called the σ-algebra *generated* by $\mathcal{F}$. Except if otherwise specified, any finite set U is equipped with its *discrete σ-algebra*, that is $\mathfrak{F}$ is simply the powerset of U.

If $(U, \mathfrak{F})$ is a measurable space, any measurable subset A is equipped with the σ-algebra $\mathfrak{F}^A$ *induced by* $\mathfrak{F}$, defined by:

$$\mathfrak{F}^A = \{B \in \mathfrak{F} \,:\, B \subseteq A\} = \{B \cap A, \ B \in \mathfrak{F}\}.$$

Probability Measures. If $(U, \mathfrak{F})$ is a measurable space, the triple $(U, \mathfrak{F}, \mathbb{P})$ is said to be a *probability space* if $\mathbb{P}$ is a nonnegative set function $\mathbb{P} : \mathfrak{F} \to \mathbb{R}$ such that $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$, and for any sequence $(A_n)_{n \geq 0}$ of pairwise disjoint measurable subsets, we have $\mathbb{P}\big(\bigcup_{n \geq 0} A_n\big) = \sum_{n \geq 0} \mathbb{P}(A_n)$. $\mathbb{P}$ is called a *probability measure*, or simply a *probability*.

If U is a finite set, equipped with the discrete σ-algebra $\mathfrak{F}$, a probability $\mathbb{P}$ is entirely determined by the values of $\mathbb{P}$ on the singletons $\mathbb{P}(\{x\})$, $x \in U$. We simply note $\mathbb{P}(x) = \mathbb{P}(\{x\})$, and we have $\sum_{x \in U} \mathbb{P}(x) = 1$. Conversely, for any nonnegative function $f : U \to \mathbb{R}$ such that $\sum_{x \in U} f(x) = 1$, there is a unique probability $\mathbb{P}$ on $(U, \mathfrak{F})$ such that $\mathbb{P}(x) = f(x)$, defined by $\mathbb{P}(A) = \sum_{x \in A} f(x)$ for $A \subseteq U$.

If $\varphi : U \to V$ is a measurable mapping from $(U, \mathfrak{F})$ to $(V, \mathfrak{G})$, and if $\mathbb{P}$ is a probability measure on $(U, \mathfrak{F})$, the following formula defines $\mathbb{Q}$ as a probability measure on $(V, \mathfrak{G})$: $\mathbb{Q}(A) = \mathbb{P}\big(\varphi^{-1}(A)\big)$ for $A \in \mathfrak{G}$. $\mathbb{Q}$ is called the *image of* $\mathbb{P}$ *under* φ, and we denote it $\mathbb{Q} = \varphi\mathbb{P}$.

Let $(U, \mathfrak{F}, \mathbb{P})$ be a probability space, and let A be a measurable subset of U such that $\mathbb{P}(A) > 0$. Define the *conditional probability given A* by:

$$\forall B \in \mathfrak{F}^A, \quad \mathbb{P}(B|A) = \frac{\mathbb{P}(B)}{\mathbb{P}(A)}.$$

If $(U, \mathfrak{F}, \mathbb{P})$ is a probability space, still following the usual terminology, we define the *(mathematical) expectation* of a real-valued nonnegative random variable f as its integral, denoted by $\mathbb{E}(f)$, so that $\mathbb{E}(f) = \int_\Omega f(\omega) \, \mathrm{d}\mathbb{P}(\omega)$. If $\mathfrak{G}$ is a sub-σ-algebra of $\mathfrak{F}$, for every nonnegative $\mathfrak{F}$-measurable real-valued function f, there exists a $\mathfrak{G}$-measurable function g such that $\mathbb{E}(fh) = \mathbb{E}(gh)$ for every nonnegative $\mathfrak{G}$-measurable real-valued function h. Function g is unique up to a set of probability zero, it is called the *conditional expectation of f given $\mathfrak{G}$*, denoted by $\mathbb{E}(f|\mathfrak{G})$.

Isomorphisms of Probability Spaces. Let $(U, \mathfrak{F}, \mathbb{P})$ and $(V, \mathfrak{G}, \mathbb{Q})$ be two probability spaces. We say that they are *isomorphic* if there are two measurable subsets $U' \subseteq U$ and $V' \subseteq V$ such that $\mathbb{P}(U') = 1$ and $\mathbb{Q}(V') = 1$, and a bi-measurable bijective mapping $\varphi : U' \to V'$ such that:

$$\varphi\mathbb{P} = \mathbb{Q}, \quad \varphi^{-1}\mathbb{Q} = \mathbb{P}.$$

In the above equations, $\mathbb{P}$ and $\mathbb{Q}$ must be understood as their restriction to the induced σ-algebras $\mathfrak{F}^{U'}$ and $\mathfrak{G}^{V'}$ respectively.

4.2 Probabilistic Event Structures

An event structure E naturally defines a measurable space as follows. Consider first, for any configuration v of E, the following nonempty subset of Ω_E:

$$\mathcal{S}(v) =_{\mathrm{def}} \{\omega \in \Omega_E \, : \, \omega \supseteq v\}.$$

$\mathcal{S}(v)$ is called the *shadow* of v. We define the *Borel σ-algebra of Ω_E* as the σ-algebra generated by the collection of shadows $\mathcal{S}(v)$, where v ranges over the finite configurations of E. This is indeed the Borel σ-algebra generated by the Scott topology on Ω (see [1, Ch.2, III-1.*1*] for details). Unless otherwise specified, Ω_E *is always equipped with the Borel σ-algebra*, and thus we simply omit it.

The following definition has already been considered in [15, 6, 1], see also the *probabilistic runs* of [14].

Definition 4.1 (probabilistic event structure, likelihood) *A* probabilistic event structure *is a pair $(E, \mathbb{P})$ where E is an event structure and $\mathbb{P}$ is a probability measure on the space Ω of maximal configurations of E.*

If $(E, \mathbb{P})$ is a probabilistic event structure, we define the likelihood *of $\mathbb{P}$ as the real-valued function $p : \mathcal{V} \to \mathbb{R}$ defined by:*

$$\forall v \in \mathcal{V}, \quad p(v) = \mathbb{P}\big(\mathcal{S}(v)\big).$$

We say that a probabilistic event structure $(E, \mathbb{P})$ is positive *if we have:*

$$\forall v \in \mathcal{V}, \quad p(v) > 0.$$

Next, consider a probabilistic event structure $(E, \mathbb{P})$ and define the following two notions:

1. **Restriction to stopping prefixes.** Let B be a stopping prefix of E, and let $\pi_B : \Omega \to \Omega_B$ be the mapping defined by $\pi_B(\omega) = \omega \cap B$, which is well defined according to Lemma 3.1. π_B is a measurable mapping. We define the probability $\mathbb{P}_B$ on Ω_B by:

$$\mathbb{P}_B = \pi_B \mathbb{P},$$

image of $\mathbb{P}$ under π_B. In particular, if B is a finite stopping prefix, $(\Omega_B, \mathbb{P}_B)$ is a finite probability space, and we have:

$$\forall v \in \Omega_B, \quad \mathbb{P}_B(v) = \mathbb{P}\big(\mathcal{S}(v)\big).$$

2. **Probabilistic future.** Let v be a finite configuration of E such that $p(v) > 0$. $\mathcal{S}(v)$ is then a measurable subset of Ω with positive probability. It is thus equipped with the conditional probability, denoted by $\mathbb{P}^v$, and defined as follows, for A a measurable subset of $\mathcal{S}(v)$:

$$\mathbb{P}^v(A) = \frac{1}{p(v)} \mathbb{P}(A).$$

Consider the bijective and bi-measurable mapping:

$$\Phi^v : \mathcal{S}(v) \to \Omega_{E^v}, \quad \omega \mapsto \omega \ominus v.$$

We still denote by $\mathbb{P}^v$ the probability on Ω_{E^v}, image of $\mathbb{P}^v$ under Φ^v. For every finite configuration v with $p(v) > 0$, we define the probabilistic event structure $(E^v, \mathbb{P}^v)$ thus obtained as the *probabilistic future of v*. The likelihood p^v of $(E^v, \mathbb{P}^v)$ is given by:

$$\forall w \in \mathcal{V}_{E^v}, \quad p^v(w) = \frac{1}{p(v)} p(v \oplus w). \tag{12}$$

4.3 Locally Randomized Event Structures

We have shown in [3] that a probabilistic event structure can be naturally defined from the new notion of *locally randomized* event structure. We recall this construction.

Definition 4.2 (locally randomized event structure) *A* locally randomized event structure *is a pair $(E, (p_x)_{x \in X})$, where X is the set of branching cells of E, and for each $x \in X$, p_x is a probability on Ω_x. Say that $(E, (p_x)_{x \in X})$ is* positive *if*

$$\forall x \in X_E, \quad \forall z \in \Omega_x, \quad p_x(z) > 0.$$

Observe that, in this definition, the probabilities p_x are finite probabilities since branching cells are finite by Theorem 3.1, point 1, and thus every Ω_x is *a fortiori* finite. Let

$\big(E, (p_x)_{x \in X}\big)$ be a locally randomized event structure. For B a finite stopping prefix of E, we set:

$$\forall \omega_B \in \Omega_B, \quad \mathbb{Q}_B(\omega_B) = \prod_{x \in \Delta(\omega_B)} p_x(\omega_B \cap x), \tag{13}$$

which is well defined since, according to Theorem 3.1, point 3, $\omega_B \cap x \in \Omega_x$. Remark that, if $B = B_0(E)$ is the max-initial stopping prefix of E, then $\mathbb{Q}_B$ coincides with the product probability on Ω_B:

$$\mathbb{Q}_B = \bigotimes_{x \in \delta_E(\emptyset)} p_x.$$

This is the probabilistic counterpart of decomposition $\Omega_B = \prod_{x \in \delta(\emptyset)} \Omega_x$, stated in Equation (8), point 2 of Theorem 3.1. The product form of probability $\mathbb{Q}_B$ manifests that "local actions" associated to initial stopping prefixes $x \in \delta(\emptyset)$ are independent in the probabilistic sense. The family $(\mathbb{Q}_B)_B$ makes E a probabilistic event structure, as expressed by the following theorem:

Theorem 4.1 (distributed product) *Let $\big(E, (p_x)_{x \in X}\big)$ be a locally randomized event structure.*

1. ***Distributed Product and Distributed Probabilities.*** *There exists a unique probabilistic event structure $(E, \mathbb{P})$ such that $\mathbb{P}_B = \mathbb{Q}_B$ for every finite stopping prefix $B \subseteq E$, where $\mathbb{P}_B$ denotes the restriction $\pi_B \mathbb{P}$ of $\mathbb{P}$ to Ω_B. The probability $\mathbb{P}$ is called the* distributed product *of the family $(p_x)_{x \in X}$, written:*

$$\mathbb{P} = \bigotimes_{x \in X}^{\mathrm{dist}} p_x.$$

For every finite R-stopped configuration v, we have:

$$p(v) = \prod_{x \in \Delta(v)} p_x(v \cap x). \tag{14}$$

In particular, the locally randomized event structure $\big(E, (p_x)_{x \in X}\big)$ is positive if and only if the probabilistic event structure $(E, \mathbb{P})$ is positive. Probabilities $\mathbb{P}$ arising from a distributed product are called distributed probabilities.

2. ***Restriction to stopping prefixes.*** *For every stopping prefix $B \subseteq E$, the restriction $\mathbb{P}_B = \pi_B \mathbb{P}$ coincides with the distributed product of the family $(p_x)_{x \in X_B}$, which is the restriction of $(p_x)_{x \in X}$ to the branching cells of B:*

$$\mathbb{P}_B = \bigotimes_{x \in X_B}^{\mathrm{dist}} p_x.$$

3. **Probabilistic future.** *Assume that $\big(E, (p_x)_{x \in X}\big)$ is positive, so that $(E, \mathbb{P})$ is a positive probabilistic event structure. Then for every finite R-stopped configuration v, the probabilistic future $(E^v, \mathbb{P}^v)$ coincides with the distributed product of the locally randomized event structure $\big(E^v, (p_x)_{x \in X^v}\big)$, obtained by restricting family $(p_x)_{x \in X}$ to set X^v of all branching cells of E^v:*

$$\forall v \in \mathcal{W}_E, \quad \mathbb{P}^v = \bigotimes_{x \in X^v}^{\mathrm{dist}} p_x \,.$$

REMARK. Point 2 is almost immediate. We mention it explicitly to underline the symmetry with point 3. Formula (14) extends (13) from stopped configurations to R-stopped configurations. Formula (14) also shows that, for confusion-free event structures, the *valuations with independence* defined in [14] are equivalently defined as functions of the form (14) associated with distributed products.

Corollary 4.1   *Let $\big(E, (p_x)_{x \in X}\big)$ be a locally randomized event structure, and let $(E, \mathbb{P})$ be the associated distributed product. For any finite and compatible R-stopped configurations v and w, if we set $u = v \cap w$, we have:*

$$p(v \cup w) = \frac{1}{p(u)}\, p(v)\, p(w).$$

In words, two compatible configurations v and w are probabilistically independent conditionally on their common past. This property expresses that "concurrency matches probabilistic independence", at the grain of branching cells [3]. The fact that branching cells are minimal with this property is discussed in [3].

On the other hand, if a distributed probability $\mathbb{P}$ on Ω_E is positive, it is shown in [3] that there is a *unique* locally randomized event structure $\big(E, (p_x)_{x \in X}\big)$ such that $\mathbb{P}$ is the distributed product of $(p_x)_{x \in X}$.

5   Markov Nets

In this section, we apply the previous probabilistic constructions to event structures arising from the unfolding of safe Petri nets. Recall that such an event structure E is labeled by the transitions of the considered net. It is therefore natural to consider the subclass of associated locally randomized event structures such that the labeling $x \mapsto p_x$, of the branching cells x of E by local probability p_x, conforms the labeling of events. That is, we require that the locally randomized event structures $\big(E, (p_x)_{x \in X}\big)$ satisfies $p_x = p_{x'}$ whenever x, x' are branching cells isomorphic when seen as labeled event structures. This leads to the notion of *Markov net*, a proper generalization of discrete Markov chains to true-concurrent systems (see §5.2 for a detailed discussion of the latter claim).

For the unfolding of a safe Petri net, equivalence classes of branching cells up to isomorphism of labeled event structures are finitely many. We call them *dynamic clusters*. We shall

argue that dynamic clusters are an appropriate concept of local state, for Petri nets. The main theorem of this section, the Law of Large Numbers (LLN) for Markov nets, supports this claim. Indeed, we show that the LLN holds and generalizes the LLN for Markov chains, provided that the set of dynamic clusters is taken as the state space.

Local finiteness of the unfolding is assumed. This is a non-trivial restriction on the net, although this class of Petri nets is strictly larger than the classes of free-choice or confusion-free nets. As was said in the Introduction, key elements of our approach remain valid in a more general setting—see Lemma 7.1 and the comment that follows.

5.1   Definition and First Properties of Markov Nets

$\mathcal{N}$ generically denotes a safe Petri net. Denote by E the canonical event structure that unfolds $\mathcal{N}$ (as recalled in §3.1). For v a finite configuration of E, we denote by $\gamma(v)$ the marking reached in $\mathcal{N}$ after the action of configuration v. If m is a reachable marking, we denote by $\mathcal{N}^m$ the safe Petri net identical to $\mathcal{N}$, except that $\mathcal{N}$ has m as initial marking. We also denote by E^m the unfolding of $\mathcal{N}^m$.

It is well known that, if v is a finite configuration of E, there is a *unique* isomorphism of labeled event structures $E^m \to E^v$ (see a proof in [1, Ch.5, I-2.5]). This makes the notation E^m coherent with our previous notation E^v for the future of configurations, so that we can write $E^v = E^{\gamma(v)}$. We can also rewrite it as follows:

$$\forall v, v' \in \mathcal{V}_E, \quad \gamma(v) = \gamma(v') \Rightarrow E^v = E^{v'} . \tag{15}$$

Finally, for m a reachable marking, we denote by γ^m the map that is defined on the poset of finite configurations of E^m and such that $\gamma^m(w)$ is the marking reached by w from marking m, for $w \in \mathcal{V}_{E^m}$.

Recall that two T-labeled event structures (E, λ) and (E', λ'), i.e., two event structures equipped with mappings $\lambda : E \to T$ and $\lambda' : E' \to T'$ are said to be *isomorphic* if there is a mapping $\Phi : E \to E'$ such that:

1. $\forall e, e' \in E, \quad e \preceq e' \iff \Phi(e) \preceq \Phi(e')$ and $e \# e' \iff \Phi(e) \# \Phi(e')$;

2. $\lambda' = \Phi \circ \lambda$.

Since the reachable markings are finitely many, the futures $E^v = E^{\gamma(v)}$ are finitely many up to isomorphism of labeled event structures. Since each set of branching cells $\delta_E(v)$ is finite, it follows then from Definition 3.5 that branching cells of E are finitely many, *up to an isomorphism of labeled event structures.*

Definition 5.1 (dynamic cluster)   *An isomorphism class of branching cells is called a* dynamic cluster *of* $\mathcal{N}$*. We denote by* Σ *the (finite) set of dynamic clusters of* $\mathcal{N}$*. Dynamic clusters are generically denoted by the boldface symbol* **s***. The equivalence class of branching cell x is denoted by $\langle x \rangle$.*

REMARK. It is shown in [4] that, if the event structure is confusion-free, branching cells can be interpreted as the events of a new event structure, called *choice structure*. The set of dynamic clusters Σ is then a finite alphabet that labels the choice structure. Under certain conditions, the labeled event structure obtained is actually itself the unfolding of a safe Petri net, called the *choice net*. The interested reader is referred to [4] for further details.

Lemma 5.1 *If x and x' are two isomorphic branching cells of the unfolding E, then the isomorphism $x \to x'$ is unique.*

Proof. Let v be a finite R-stopped configuration of E such that $x \in \delta_E(v)$. Let m be the set of places b of $\mathcal{N}$ such that there is an event $e \in x$, minimal in x, and with $b \, F \, \lambda(e)$; i.e., b is in the preset of $\lambda(e)$. Consider the subnet $\mathcal{N}'$ of $\mathcal{N}$ with same sets of places and transitions, but with initial marking m'. Then the unfolding F of $\mathcal{N}'$ has a unique initial stopping prefix y, isomorphic to x. By the uniqueness property of unfoldings, the isomorphism $x \to y$ is unique.

Symmetrically, there is a unique isomorphism $x' \to y$. Hence the isomorphism $x \to x'$ is unique, which completes the proof of the lemma. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

As a consequence, for every dynamic cluster $\mathbf{s}$ and every pair of branching cells $x, x' \in \mathbf{s}$, there exists a *canonical* bijection $\phi_{x,x'} : \Omega_x \to \Omega_{x'}$, namely, the bijection induced by the unique isomorphism $x \to x'$. Hence, we can consistently consider the space $\Omega_\mathbf{s}$. In particular, we may define a real-valued function $g : \Omega_\mathbf{s} \to \mathbb{R}$, by means of a family of functions $(g_x)_{x \in \mathbf{s}}$, $g_x : \Omega_x \to \mathbb{R}$, such that $g_x = g_{x'} \circ \phi_{x,x'}$ for all $x, x' \in \mathbf{s}$. This is what is meant in the following definition, for the probability $p_\mathbf{s}$ on $\Omega_\mathbf{s}$ seen as a function on $\Omega_\mathbf{s}$.

Definition 5.2 (Markov net, local transition probabilities) *A Markov net is a pair $\big(\mathcal{N}, (p_\mathbf{s})_{\mathbf{s} \in \Sigma}\big)$, where $\mathcal{N}$ is a finite safe Petri net with locally finite unfolding, and $p_\mathbf{s}$ is a probability on the finite set $\Omega_\mathbf{s}$ for every $\mathbf{s} \in \Sigma$. Probability $p_\mathbf{s}$ is called the* local transition probability *attached to $\mathbf{s} \in \Sigma$. We assume moreover the following:*

$$\forall \mathbf{s} \in \Sigma, \quad \forall z \in \Omega_\mathbf{s}, \quad p_\mathbf{s}(z) > 0.$$

A Markov net $\big(\mathcal{N}, (p_\mathbf{s})_{\mathbf{s} \in \Sigma}\big)$ induces a locally randomized event structure $(E, (p_x)_{x \in X})$ by setting $p_x = p_{\langle x \rangle}$ for every branching cell $x \in X_E$; in turn, the distributed product $\mathbb{P}$ of the family $(p_x)_{x \in X}$ defines a probabilistic event structure $(E, \mathbb{P})$. Furthermore, this probabilistic event structure is *positive*, according to point 3 in Theorem 4.1.

Note that, if net $\mathcal{N}$ is composed of two disjoint, and thus non interacting nets $\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2$, then the two components $\mathcal{N}_i, i \in \{1, 2\}$ are independent in the probabilistic sense, i.e., $\mathbb{P} = \mathbb{P}_1 \otimes \mathbb{P}_2$: once again, "probabilistic independence matches concurrency".

Theorem 5.1 (Homogeneity) *Let $\big(\mathcal{N}, (p_\mathbf{s})_{\mathbf{s} \in \Sigma}\big)$ be a Markov net, and let $\mathbb{P}$ be the associated distributed probability on Ω. For v a finite R-stopped configuration of E, we use the notations $\gamma(v)$ and $\mathcal{N}^{\gamma(v)}$ introduced above, and we let $\Sigma^{\gamma(v)}$ denote the set of dynamic clusters of $\mathcal{N}^{\gamma(v)}$.*

For any $v \in \mathcal{W}$, *the probabilistic future* $(E^v, \mathbb{P}^v)$ *is associated with Markov net* $(\mathcal{N}^{\gamma(v)}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma^{\gamma(v)}})$. *Moreover we have:*

$$\forall v, v' \in \mathcal{W}, \quad \gamma(v) = \gamma(v') \Rightarrow \mathbb{P}^v = \mathbb{P}^{v'}. \tag{16}$$

Proof. According to Theorem 4.1, point 3, we have, for $v \in \mathcal{W}_E$:

$$\mathbb{P} = \bigotimes_{x \in X}^{\text{dist}} p_x, \quad \mathbb{P}^v = \bigotimes_{x \in X^v}^{\text{dist}} p_x,$$

where X^v is the set of branching cells of E^v. Therefore, $\mathbb{P}^v$ is indeed the distributed product associated with the Markov net $(\mathcal{N}^{\gamma(v)}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma^{\gamma(v)}})$.

For $v, v' \in \mathcal{W}$ with $\gamma(v) = \gamma(v')$, let $\varphi : E^v \rightarrow E^{v'}$ denote the unique isomorphism of labeled event structures. We denote by $\Phi : X^v \rightarrow X^{v'}$ the induced bijection between the two sets of branching cells. We have:

$$\mathbb{P}^v = \bigotimes_{x \in X^v}^{\text{dist}} p_x = \bigotimes_{y \in X^{v'}}^{\text{dist}} p_{\Phi^{-1}(y)}, \tag{17}$$

by the change of variable $x = \Phi^{-1}(y)$, which is allowed since Φ is bijective. For every branching cell $x \in X^v$, the restriction $\varphi|_x : x \rightarrow \Phi(x)$, with $x \in X^v$, does not change the class of branching cell. Hence, up to isomorphism of branching cells, $\varphi|_x$ is the identity map. Therefore $p_{\Phi^{-1}(y)} = p_y$ for every $y \in X^{v'}$. Hence, from (17), we get:

$$\mathbb{P}^v = \bigotimes_{y \in X^{v'}}^{\text{dist}} p_y = \mathbb{P}^{v'}.$$

This completes the proof of the theorem. □

Equation (16) expresses that the memory of Markov nets is entirely summarized by the current marking: the probabilistic future of a $v \in \mathcal{W}$ only depends on the final marking $\gamma(v)$, and not on the entire history v. It is the probabilistic counterpart of Equation (15). In the setting of Markov chains, this is equivalent to the time-invariance property of the transition matrix, characteristic of *homogeneous Markov chains* (see e.g., [7, 13]). An important consequence of homogeneity, for Markov chains, is the so-called *strong Markov property*; see a generalization for Markov nets in [2].

It will be convenient to use the following terminology and notation:

Definition 5.3 (recursively stopped marking) *We say that a marking m is recursively stopped if there is a finite recursively stopped configuration v such that $\gamma(v) = m$. The set of recursively stopped markings is denoted by $\mathbf{M}_{\text{rs}}$.*

Thanks to property (16), we may define for each $m \in \mathbf{M}_{\text{rs}}$, the probabilistic event structure $(E^m, \mathbb{P}^m)$, by defining $\mathbb{P}^m$ as the probability $\mathbb{P}^v$ on Ω_{E^m}, whenever v is a finite R-stopped configuration such that $\gamma(v) = m$. Moreover, Theorem 5.1 says that $\mathbb{P}^m$ is the probability associated with the Markov net $(\mathcal{N}^m, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma^m})$.

5.2   Markov Nets as a Generalization of Markov Chains

For further referencing it will be useful to see how Markov nets are a generalization of Markov chains. To this end, we construct for each Markov chain a canonical Markov net that simulates the chain.

Let $(X_n)_{n \geq 0}$ be a Markov chain defined on a finite state space P, with transition matrix $(\rho_{i,j})_{(i,j) \in P \times P}$ and initial state s_0. Consider the following safe Petri net $\mathcal{N} = (P, T, F, m_0)$. The set of places P coincides with the set of states of the chain, and $m_0 = \{s_0\}$. The set T of transitions is defined by this rule: $T \subseteq P \times P$, with:

$$\forall (i,j) \in P \times P, \quad (i,j) \in T \iff \rho_{i,j} > 0.$$

Then the flow relation $F \subseteq (P \times T) \cup (T \times P)$ is naturally defined by:

$$\forall (i,j) \in P \times P, \quad (i,j) \in T \Rightarrow i \, F \, (i,j) \quad \text{and} \quad (i,j) \, F \, j.$$

Each reachable marking of $\mathcal{N}$ is a singleton, and there is a one-to-one and onto correspondence between paths of the chain and firing sequences of the net. The unfolding of $\mathcal{N}$ is a tree. It is thus clear that the set Σ of dynamic clusters of $\mathcal{N}$ is given by $\Sigma = \{\mathbf{s}_i, \, i \in P\}$, with:

$$\forall i \in P, \quad \mathbf{s}_i = \{(i,j), \, j \in P \text{ such that } (i,j) \in T\}.$$

Hence, for Markov chains, *dynamic clusters identify with states* of the chain.

To define a Markov net $\big(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma}\big)$, it remains only to define the family of local transition probabilities $(p_{\mathbf{s}})_{\mathbf{s} \in \Sigma}$. This is done as follows, using the fact that maximal configurations of a cluster $\mathbf{s}_i$ identify with the transitions (i,j) of $\mathbf{s}_i$:

$$\forall i \in P, \quad \forall j \in P, \quad (i,j) \in \mathbf{s}_i \Rightarrow p_{\mathbf{s}_i}(i,j) = \rho_{i,j}.$$

That is, *local transition probabilities are given by the rows of the transition matrix*. Remark that we have by construction $p_{\mathbf{s}}(z) > 0$, for all $\mathbf{s} \in \Sigma$ and $z \in \Omega_{\mathbf{s}}$.

Finally, we must prove that, if $(\Omega, \mathfrak{F}, \mathbb{P})$ is the probability space associated with the runs of Markov net $\big(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma}\big)$, and if $(\Xi, \mathfrak{G}, \mathbb{Q})$ is the probability space associated with the runs of the Markov chain $(X_n)_{n \geq 0}$, there is an isomorphism of probability spaces:

$$(\Omega, \mathfrak{F}, \mathbb{P}) \to (\Xi, \mathfrak{G}, \mathbb{Q}).$$

This can be done by hand, by checking that formula (14) that defines the likelihood function for the Markov net coincides with the equivalent for the Markov chain. But this is also a consequence of the more general result stated in Lemma 7.1 below, §7.1.

5.3   Recurrent Nets

Recurrent Markov chains are chains in which almost every trajectory returns infinitely often to the initial state. As a generalization, recurrent Markov nets will be nets in which almost

every firing sequence returns infinitely often to the initial marking. The above statement calls for distinguishing between two types of returns: "global" returns, where all tokens must leave the initial marking before returning to it, and "local" returns, where some "minimal" amount of tokens move. Global return is investigated next; the study of local return is postponed to §7.3.

Definition 5.4 (return operator) *Let m_0 be the initial marking of $\mathcal{N}$. We define the* return operator R *on* Ω *as follows: for every* $\omega \in \Omega$,

$$R(\omega) = \min\{v \in \mathcal{W}_E \,:\, v \subseteq \omega, \, \gamma(v) = m_0, \, \mathrm{Min}(E) \cap \mathrm{Min}(E^v) = \emptyset\}, \qquad (18)$$

with the convention that $R(\omega) = \omega$ if the set of such v is empty.

The condition $\min(E) \cap \mathrm{Min}(E^v) = \emptyset$ says that configuration v has moved all the tokens in the net. In other words, we do not allow the return to leave some part of the marking untouched, while acting on the other tokens.

It is well known that, if v, v' are two compatible configurations such that $\gamma(v) = \gamma(v') = m_0$, then $\gamma(v \cap v') = m_0$. Furthermore, the intersection of compatible R-stopped configurations is still R-stopped thanks to Theorem 3.1, point 5. Finally, it is also readily checked that the intersection of configurations satisfying the condition $\mathrm{Min}(E) \cap \mathrm{Min}(E^v) = \emptyset$ also satisfies this condition. Therefore, if the set of configurations in the right member of (18) is nonempty, the minimum is well defined and is finite.

The study of properties of R is postponed to §7.2. For the moment, we define the *successive returns* by:

$$R_0 = \emptyset, \qquad \forall n > 0, \quad R_n(\omega) = R_{n-1}(\omega) \oplus R\big(\omega \setminus R_{n-1}(\omega)\big),$$

with the convention that $R_n(\omega) = \omega$ if $R_{n-1}(\omega) = \omega$. We mention without proof that the successive returns $(R_n)_{n \geq 0}$ are all finite with probability either 0 or 1 [2], which extends to nets a classical result for Markov chains.

Definition 5.5 (recurrent nets) *We say that a Markov net $\big(\mathcal{N}, (p_\mathbf{s})_{\mathbf{s} \in \Sigma}\big)$ is* recurrent *if the successive returns satisfy $R_n(\omega) \neq \omega$ for all $n \geq 0$, with probability 1.*

If the considered net is a simulated Markov chain, as described in §5.2, then this definition of recurrence reduces to the classical notion of recurrence for Markov chains [7, 13], and the R_n are the nth returns to the initial state of the chain. Also, the following lemma extends a classical result for Markov chains [2]:

Lemma 5.2 *If $\big(\mathcal{N}, (p_\mathbf{s})_{\mathbf{s} \in \Sigma}\big)$ is a recurrent Markov net, then $(\mathcal{N}^m, (p_\mathbf{s})_{\mathbf{s} \in \Sigma^m})$ is recurrent for every $m \in \mathbf{M}_{\mathrm{rs}}$, and $\Sigma^m = \Sigma$.*

6 The Law of Large Numbers (LLN)

For our study of the Law of Large Numbers, we focus on recurrent nets.

6.1 What is the proper notion of LLN, for Markov nets?

For finite recurrent Markov chains, the LLN states as follows. Let Σ be the finite state space of a Markov chain $(X_k)_{k\geq 1}$, and let $f : \Sigma \to \mathbb{R}$ be a test function. The sums $S_n(f) = \sum_{k=1}^n f(X_k)$ are called *ergodic sums*, and the LLN studies the limit, for $n \to \infty$, of the *ergodic means*: $M_n(f) = \frac{1}{n} S_n(f)$. In extending the LLN to Markov net $\mathcal{N}$, we are faced with two difficulties:

1. What is the proper concept of state? What are the associated ergodic sums?

2. What replaces counter n, since time is not totally ordered?

Corresponding answers are:

1. The set Σ of dynamic clusters of $\mathcal{N}$ is taken as the state space—see the discussion of §5.2.

2. For v a R-stopped configuration, the number of branching cells contained in $\Delta(v)$ is taken as the "duration" of v.

Since we consider dynamic clusters as our state space, test functions are simply functions defined on Σ:

Definition 6.1 (state functions) *If Σ denotes the set of dynamic clusters of a net $\mathcal{N}$, we call* state function *any real-valued function $f : \Sigma \to \mathbb{R}$.*

State function form a vector space of finite dimension ($=\text{Card}(\Sigma)$). The *concurrent ergodic sums* associated with a state function $f : \Sigma \to \mathbb{R}$ are defined as the function $\langle f, \cdot \rangle$:

$$\langle f, \cdot \rangle : \mathcal{W} \to \mathbb{R}, \qquad \forall v \in \mathcal{W}, \quad \langle f, v \rangle = \sum_{x \in \Delta(v)} f(x),$$

where we recall that $\langle x \rangle$ denotes the dynamic cluster defined by the class of branching cell x. The scalar product notation is justified since $\langle f, v \rangle$ is linear w.r.t. its left argument for the usual addition of functions, and additive w.r.t. its right argument for the concatenation $\oplus$ of configurations.

EXAMPLE. Let **1** be the unit state function, defined by:

$$\forall \mathbf{s} \in \Sigma, \quad \mathbf{1}(\mathbf{s}) = 1. \tag{19}$$

Then $\langle \mathbf{1}, v \rangle$ counts the number of branching cells contained in $\Delta(v)$. This example will be of repeated use in the sequel.

The *concurrent ergodic means* $M(f, \cdot) : \mathcal{W} \to \mathbb{R}$ associated with a state function f are defined as the following ratios:

$$\forall v \in \mathcal{W}, \quad M(f, v) = \frac{1}{\langle \mathbf{1}, v \rangle} \langle f, v \rangle.$$

The LLN is concerned by the limit:

$$\lim_{v \subseteq \omega, v \to \omega} M(f, v),$$

where this limit is meant in a sense we shall make precise. The following notion of *stopping operator* will be central in this respect. Stopping operators generalize for concurrent systems the classical notion of stopping times [13, 7] for sequential stochastic processes in discrete time; see [2] for a detailed discussion.

Definition 6.2 (stopping operator, regular sequences) *A measurable mapping* $V : \Omega \to \mathcal{W}$, *satisfying* $V(\omega) \subseteq \omega$ *for all* $\omega \in \Omega$, *is called a* stopping operator *if for all* $\omega, \omega' \in \Omega$, *we have:*

$$\forall \omega, \omega' \in \Omega, \quad \omega' \supseteq V(\omega) \Rightarrow V(\omega') = V(\omega). \tag{20}$$

Say that a sequence $(V_n)_{n \geq 1}$ *of stopping operators is* regular *if the following properties are satisfied:*

1. $V_n \subseteq V_{n+1}$ *for all* n, *and* $\bigcup_n V_n(\omega) = \omega$, *with probability 1;*

2. There exists a constant $k > 0$ *such that:*

$$\forall n \geq 1, \quad \langle \mathbf{1}, V_n(\omega) \rangle \geq k \times n, \quad \textit{with probability 1,}$$

where $\mathbf{1}$ *is the unit state function defined in* (19).

EXAMPLE. *Stopping prefixes as stopping operators.* As an example, let B be a stopping prefix of E. Then the map $V_B : \Omega \to \Omega_B$, $\omega \mapsto \omega \cap B$, is a stopping operator. Indeed, $V_B(\omega)$ is R-stopped, and $V_B(\omega) \subseteq \omega$. Finally, let $\omega, \omega' \in \Omega$ such that $\omega' \supseteq V_B(\omega)$. Then $\omega' \cap B \supseteq \omega \cap B$. But, since $\omega \cap B$ is maximal in B, it implies that $\omega' \cap B = \omega \cap B$, which is exactly $V_B(\omega') = V_B(\omega)$. Hence V_B is indeed a stopping operator, as announced.

Not all stopping operators may be represented by this way. However, if the net arises from a Markov chain as in §5.2, then every stopping operator can be represented by a stopping prefix.

We use stopping operators to express the notion of convergence as follows:

Definition 6.3 (convergence of ergodic means) *For* f *a state function, we say that the ergodic means* $M(f, \cdot)$ *converge to a function* $\mu : \Omega \to \mathbb{R}$ *if for every regular sequence* $(V_n)_{n \geq 1}$ *of stopping operators,*

$$\lim_{n \to \infty} M(f, V_n(\omega)) = \mu(\omega), \quad \textit{with probability 1.} \tag{21}$$

The important point of this definition is that the limit μ does not depend on the regular sequence of stopping operators. It is thus intrinsic to f and $\mathbb{P}$.

REMARK. *The need for synchronization.*   Concurrency prevents (21) from holding for general recurrent Markov nets, as the following example shows. Assume that $\mathcal{N}$ decomposes as a disjoint union $\mathcal{N} = \mathcal{N}^1 \cup \mathcal{N}^2$, i.e., the two components $\mathcal{N}^1$ and $\mathcal{N}^2$ do not interact at all. The unfolding E of net $\mathcal{N}$ is the union of the unfoldings E^1 and E^2 of nets $\mathcal{N}^1$ and $\mathcal{N}^2$ respectively, and the set of dynamic clusters Σ of $\mathcal{N}$ is the union of the sets Σ^1 and Σ^2 of dynamic clusters of $\mathcal{N}^1$ and $\mathcal{N}^2$. We assume that both nets $\mathcal{N}^1$ and $\mathcal{N}^2$ are recurrent. Let f be the state function defined by:

$$\forall \mathbf{s} \in \Sigma, \quad f(\mathbf{s}) = \begin{cases} 1, & \text{if } \mathbf{s} \in \Sigma^1, \\ 0, & \text{if } \mathbf{s} \in \Sigma^2. \end{cases}$$

The associated ergodic sums $\langle f, v \rangle$ count the number of occurrences of dynamic clusters from Σ^1 in v. As shown by Lemma 6.1 below, there are two regular sequences $V^i = (V_n^i)_{n \geq 0}$, of stopping operators of $\mathcal{N}^i$ respectively, for $i = 1, 2$, such that for some constants $k, K > 0$:

$$i = 1, 2, \quad \forall n > 1, \quad k \times n \leq \langle \mathbf{1}, V_n^i \rangle \leq K \times n.$$

Define, for $n > 1$, $V_n = V_n^1 \cup V_{[\log n]}^2$, where $[\log n]$ denotes the greatest integer less than $\log n$. Define symmetrically $W_n = V_{[\log n]}^1 \cup V_n^2$. Then $(V_n)_{n>1}$ and $(W_n)_{n>1}$ are two regular sequences of stopping operators of E. We compute the ergodic means of f along each of the sequences $(V_n)_{n>1}$ and $(W_n)_{n>1}$. We have:

$$\begin{aligned} M(f, V_n) = M(f, V_n^1 \cup V_{[\log(n)]}^2) &= \frac{\langle \mathbf{1}, V_n^1 \rangle}{\langle \mathbf{1}, V_n^1 \rangle + \langle \mathbf{1}, V_{[\log(n)]}^2 \rangle} \\ &\geq \frac{kn}{Kn + K[\log(n)]} \longrightarrow_{n \to \infty} \frac{k}{K} > 0. \end{aligned}$$

On the other hand:

$$\begin{aligned} M(f, W_n) = M(f, V_{[\log(n)]}^1 \cup V_n^2) &= \frac{\langle \mathbf{1}, V_{[\log(n)]}^1 \rangle}{\langle \mathbf{1}, V_n^1 \rangle + \langle \mathbf{1}, V_{[\log(n)]}^2 \rangle} \\ &\leq \frac{K[\log(n)]}{kn + k[\log(n)]} \longrightarrow_{n \to \infty} 0. \end{aligned}$$

Hence, the limit of the ergodic means depends on the particular regular sequence of stopping operators chosen. Thus, the convergence of ergodic means does not hold in the sense of Definition 6.3.

Clearly, the concurrency properties of $\mathcal{N}$ is the very cause of this difficulty in the example; more precisely, the total absence of synchronization brings this behavior. We show below "how much synchronization" is needed for the LLN to be valid.

The existence of regular sequences of stopping operators results from the following lemma. Recall that we have defined in §3.6 the normal decomposition $(V_n(\omega), Z_{n+1}(\omega))_{n \geq 0}$ of a maximal configuration ω. Since V_n and Z_n are (measurable) mappings defined on Ω, we now see this sequence as a sequence of operators, as in Definition 6.2.

Lemma 6.1 *For every $n \geq 0$, the mapping $\omega \mapsto V_n(\omega)$ resulting from the normal decomposition of maximal configuration ω is a stopping operator. There is a constant $K > 0$ such that:*

$$\forall n \geq 0, \quad \langle \mathbf{1}, V_n \rangle \leq K \times n. \tag{22}$$

If the net is recurrent, then $(V_n)_{n \geq 0}$ is a regular sequence of stopping operators.

Proof. By construction, $V_n(\omega) \subseteq \omega$ for all $\omega \in \Omega$. We prove the other point of the definition of stopping operators, stated by Equation (20) in Definition 6.2, by induction on $n \geq 0$. This is trivial for $n = 0$; assume that Equation (20) holds for V_n until $n \geq 0$. Let $\omega, \omega' \in \Omega$ such that $\omega' \supseteq V_{n+1}(\omega)$. Then in particular $\omega' \supseteq V_n(\omega)$, so that $V_n(\omega') = V_n(\omega)$ by the induction hypothesis. Put $v = V_n(\omega)$, $\xi = \omega \ominus v$ and $\xi' = \omega' \ominus v$, so that $\xi, \xi' \in \Omega_{E^v}$. We have:

$$Z_{n+1}(\omega') = \xi' \cap B_0(E^v) \supseteq \xi \cap B_0(E^v).$$

Since $\cdot \cap B_0(E^v)$ is a stopping operator of Ω_{E^v}, since it is defined by means of the stopping prefix $B_0(E^v)$, it follows that $Z_{n+1}(\omega') = \xi' \cap B_0(E^v) = \xi \cap B_0(E^v) = Z_{n+1}(\omega)$. Finally, $V_{n+1}(\omega') = V_n(\omega') \oplus Z_{n+1}(\omega') = V_{n+1}(\omega)$, which completes the proof of (20). This shows that every V_n is a stopping operator.

We now show Equation (22). According to Equation (10) in Theorem 5, we have:

$$\forall n > 0, \qquad \Delta(V_n) = \Delta(V_{n-1}) \cup \Delta(Z_n), \quad \Delta(V_{n-1}) \cap \Delta(Z_n) = \emptyset,$$

with Δ the covering map. Therefore:

$$\forall n \geq 0, \quad \langle \mathbf{1}, V_n \rangle = \mathrm{Card}\big(\Delta(V_n)\big) = \sum_{i=1}^{n} \mathrm{Card}\big(\Delta(Z_i)\big). \tag{23}$$

As observed in § 3.5, since Z_i has the form $Z_i = \bigcup_{x \in \delta_E(V_{i-1})} Z_i \cap x$ with $Z_i \cap x \in \Omega_x$ for $x \in \delta_E(V_{i-1})$, the covering $\Delta(Z_i)$ is given by $\Delta(Z_i) = \delta_E(V_{i-1})$. Therefore, according to point 2b of Theorem 3.1, there is a constant K such that $\mathrm{Card}(\Delta(Z_i)) \leq K$. Hence it follows from (23) that:

$$\forall n \geq 0, \quad \langle \mathbf{1}, V_n \rangle \leq K \times n,$$

which is (22).

Now we show that, if $\mathcal{N}$ is recurrent, $(V_n)_{n \geq 0}$ is a regular sequence of stopping operators. It is clear that $V_n \subseteq V_{n+1}$ for all $n \geq 0$. Moreover, according to Theorem 3.2, $\bigcup_{n \geq 0} V_n(\omega) = \omega$. Hence point 1 in Definition 6.2 is satisfied.

Since $\mathcal{N}$ is recurrent (Definition 5.5), with probability 1, ω is infinite since it contains infinitely many returns to the initial marking. Therefore, for each $i \geq 0$, V_i is not maximal, and thus $B_0(E^{V_i}) \neq \emptyset$. In particular, $\mathrm{Card}\big(\Delta(Z_{i+1})\big) \geq 1$. Hence, using Equation (23), we obtain $\langle \mathbf{1}, V_n \rangle \geq n$, for all $n \geq 0$. Hence $(V_n)_{n \geq 0}$ satisfies the definition of a regular sequence of stopping operators with $k = 1$. This completes the proof of the lemma. □

6.2 A Probabilistic Measure of Synchronization and the LLN

If, in an execution $\omega \in \Omega$, we block a token in some place b, we measure the "loss of synchronization" of the system by counting the number of branching cells that can be traversed *without moving the blocked token.* This length defines an integer random variable. A reasonable assumption is that this random variable has finite mean. We detail this definition below.

Let m be a reachable marking of a Markov net $\big(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma}\big)$, with $m \in \mathbf{M}_{\mathrm{rs}}$, and let b be a place of m. Denote by $\lambda(e)$ the transition of $\mathcal{N}$ that labels an event $e \in E$. Consider an element $\omega \in \Omega_{E^m}$, and define:

$$K_m(b, \omega) = \sup\{v \in \mathcal{W}_{E^m} \ : \ v \subseteq \omega, \, \forall e \in v, \, \lambda(e) \notin b^{\bullet}\},$$

where $b^{\bullet}$ denotes the postset of b (i.e., the set of transitions t in $\mathcal{N}$ directly after b). $K_m(b, \omega)$ describes the maximal R-stopped sub-configuration that ω allows without using the token in b. Next, define the integer $L_m(b, \omega)$ by:

$$\forall \omega \in \Omega_{E^m}, \quad L_m(b, \omega) = \begin{cases} \infty, & \text{if } K_m(b, \omega) \text{ is infinite,} \\ \langle \mathbf{1}, K_m(b, \omega) \rangle, & \text{otherwise.} \end{cases}$$

Hence $L_m(b, \omega)$ counts the number of branching cells of $K_m(b, \omega)$. It is thus a measure of non-synchronization at ω: the larger $L_m(b, \omega)$, the more sub-processes of ω can progress without synchronizing with b. For each recursively stopped marking m, and each $b \in m$, $L_m(b, \cdot)$ is now an integer random variable $\Omega_{E^m} \to \mathbb{N} \cup \{\infty\}$. Its integral is thus well-defined, although it may be infinite. We shall thus consider the following definition:

Definition 6.4 (integrable concurrency height) *Say that Markov net $\big(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma}\big)$ has integrable concurrency height if for each $m \in \mathbf{M}_{\mathrm{rs}}$, and for each place $b \in m$:*

$$\mathbb{E}^m\big(L_m(b, \cdot)\big) < \infty,$$

where $\mathbb{E}^m$ denotes the mathematical expectation under probability $\mathbb{P}^m$.

EXAMPLES. Clearly, the example given above of a net consisting of the union $\mathcal{N}^1 \cup \mathcal{N}^2$ of two non-interacting and recurrent nets has not integrable concurrency height. In contrast, if a Markov net arises from of Markov chain through the construction of §5.2, then the integrable concurrency height condition is automatically fulfilled, since the variables $L_m(b, \cdot)$ identically vanish. On the other hand, the two nets depicted in Figure 9 have integrable concurrency height.

Theorem 6.1 (Law of Large Numbers) *Let $(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma})$ be a Markov net, that we assume recurrent and with integrable concurrency height. Then:*

1. For any state function f, the ergodic means $M(f, \cdot)$ converge in the sense of Definition 6.3 to a function $\mu(f, \cdot) : \Omega \to \mathbb{R}$.

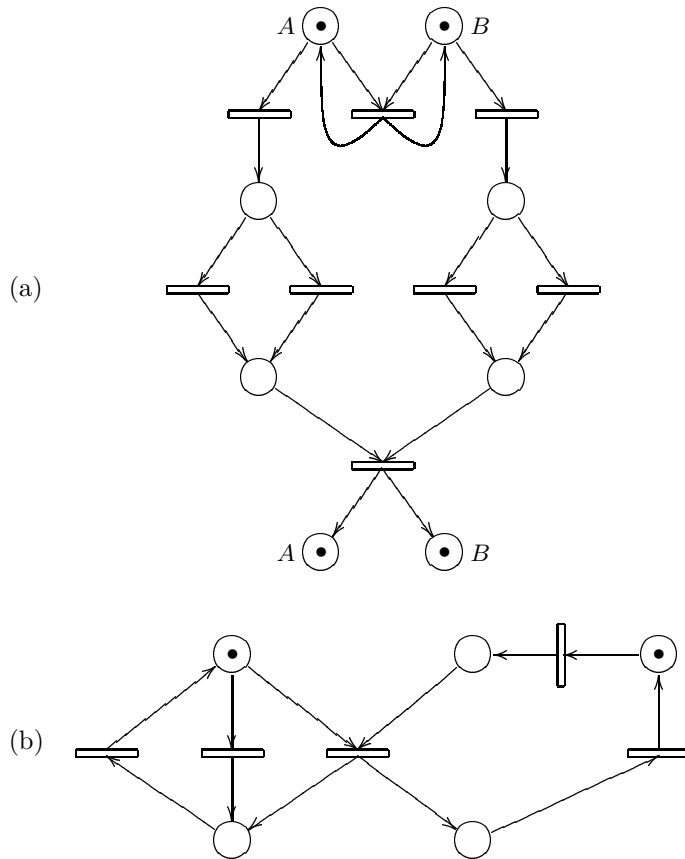Figure 9: *Two recurrent Markov nets with integrable concurrency height. Places with same labels (A and B) are identified. In (a), the random variables $L_m(b, \cdot)$ are bounded. This condition is too restrictive in general, as shown by the example (b), where the random variables $L_m(b, \cdot)$ are integrable without being bounded.*

2. *Except possibly on a set of zero probability, $\mu(f, \cdot) = \mu(f)$ is constant. There exists a family of numbers $(\alpha(\mathbf{s}))_{\mathbf{s} \in \Sigma}$ such that, for every state function $f : \Sigma \to \mathbb{R}$, the constant $\mu(f)$ is given by:*

$$\mu(f) = \sum_{\mathbf{s} \in \Sigma} \alpha(\mathbf{s}) f(\mathbf{s}). \tag{24}$$

3. *The coefficients $\alpha(\mathbf{s})$ satisfy:*

$$\forall \mathbf{s} \in \Sigma, \quad 0 < \alpha(\mathbf{s}) \le 1, \qquad and \qquad \sum_{\mathbf{s} \in \Sigma} \alpha(\mathbf{s}) = 1.$$

Hence $(\alpha(\mathbf{s}))_{\mathbf{s} \in \Sigma}$ defines a probability on Σ, we call it the stationary measure *of the Markov net.*

The proof of the theorem is the topic of §7. A closed look at the proof shows that it is enough, for Theorem 6.1 to hold, that the probability $\mathbb{P}$ satisfies the homogeneity condition; the stronger property that $\mathbb{P}$ is distributed is not required.

6.3   Interpretation of the Stationary Measure

We give an interpretation of the coefficients $\alpha(\mathbf{s})$ mentioned in Theorem 6.1. For every $\mathbf{s} \in \Sigma$, the coefficient $\alpha(\mathbf{s})$ shall be considered as the asymptotic rate of occurrence of $\mathbf{s}$ in a typical execution $\omega \in \Omega$. Indeed, consider, for $\mathbf{s}_0 \in \Sigma$, the state function $\mathbf{1}_{\mathbf{s}_0}$ defined by:

$$\forall \mathbf{s} \in \Sigma, \quad \mathbf{1}_{\mathbf{s}_0}(\mathbf{s}) = \begin{cases} 1, & \text{if } \mathbf{s} = \mathbf{s}_0, \\ 0, & \text{otherwise.} \end{cases} \tag{25}$$

The ergodic sums $\langle \mathbf{1}_{\mathbf{s}_0}, v \rangle$ count, for $v \in \mathcal{W}$, the number of occurrences of $\mathbf{s}_0$ in v. Applying the LLN to $\mathbf{1}_{\mathbf{s}_0}$, we get that, for every regular sequence of stopping operators $(V_n)_{n \ge 0}$, and with probability 1:

$$\lim_{n \to \infty} \frac{\text{number of occurrences of } \mathbf{s}_0 \text{ in } V_n}{\text{number of branching cells in } V_n} = \lim_{n \to \infty} \frac{\langle \mathbf{1}_{\mathbf{s}_0}, V_n \rangle}{\langle \mathbf{1}, V_n \rangle}$$
$$= \lim_{n \to \infty} M(\mathbf{1}_{\mathbf{s}_0}, V_n) = \mu(\mathbf{1}_{\mathbf{s}_0}) = \alpha(\mathbf{s}_0). \tag{26}$$

The expression (26) shows that $\alpha(\mathbf{s}_0)$ is the asymptotic occurrence rate of $\mathbf{s}_0$ in a typical ω, whatever the regular sequence of stopping operators V_n is. This justifies the name of density coefficient.

If the net arises from a recurrent Markov chain through the construction of §5.2, then the integrable concurrency height is satisfied, as already seen, and thus the theorem applies. The stationary measure of the net coincides with the stationary measure of the chain, in the usual sense. This again reveals that dynamic clusters play the role of local states for concurrent systems.

We now examine an example that reveals that the LLN stated in Theorem 6.1 is not fine enough to describe the asymptotic behavior of Markov nets. We will thus need to refine it.

Let $\big(\mathcal{N}, (p_\mathbf{s})_{\mathbf{s} \in \Sigma}\big)$ be a Markov net, recurrent and with integrable concurrency height as in Theorem 6.1, with unfolding (E, λ). Pick t and t' two transitions of $\mathcal{N}$. For v a finite R-stopped configuration of E, let $r_{t,t'}(v)$ be the ratio of occurrences of t and t' in v. That is:

$$r_{t,t'}(v) = \frac{\mathrm{Card}\{e \in v \,:\, \lambda(e) = t\}}{\mathrm{Card}\{e \in v \,:\, \lambda(e) = t'\}}. \tag{27}$$

We would like to know if this ratio has a limit when v grows to some $\omega \in \Omega_E$. This would tell us how much, asymptotically, transition t fires as compared to t'. However, the ratio $r_{t,t'}(v)$ cannot be expressed as a ratio of the form $\langle f, v \rangle / \langle f', v \rangle$, where f and f' would be state functions. Indeed, the occurrence of transition t and t' in an element $z \in \Omega_\mathbf{s}$, with $\mathbf{s}$ a dynamic cluster, depends on z, and not on $\mathbf{s}$ only. Hence state functions are not fine enough to evaluate quantities of the type $r_{t,t'}(v)$. This is the reason why we introduce *extended state functions*.

6.4 Extended state functions and the Extended LLN

We begin with the definition.

Definition 6.5 (extended state function) *We call* extended state function *a finite family $f = \big(f(\mathbf{s}, \cdot)\big)_{\mathbf{s} \in \Sigma}$ of real-valued functions $f(\mathbf{s}, \cdot) : \Omega_\mathbf{s} \to \mathbb{R}$, where $\mathbf{s} \in \Sigma$.*

Extended state functions extend state functions: indeed, a state function $f : \Sigma \to \mathbb{R}$ is simply an extended state function that is constant on every $\Omega_\mathbf{s}$. That is, $f(\mathbf{s}, z) = f(\mathbf{s})$ for all $\mathbf{s} \in \Sigma$ and $z \in \Omega_\mathbf{s}$. In particular, we keep the notation $\mathbf{1}$ to denote the unit extended state function, defined by $\mathbf{1}(\mathbf{s}, z) = 1$ for all $\mathbf{s} \in \Sigma$ and $z \in \Omega_\mathbf{s}$.

Ergodic sums and means of extended state functions are defined in a way that extends the definition of ergodic sums and means for state functions. The ergodic sum of an extended state function f along a finite R-stopped configuration v is defined by:

$$\langle f, v \rangle = \sum_{x \in \Delta(v)} f(\langle x \rangle, v \cap x),$$

and the ergodic means are defined by:

$$\forall v \in \mathcal{W}_E, \quad M(f, v) = \frac{\langle f, v \rangle}{\langle \mathbf{1}, v \rangle}.$$

If f is an extended state function, we define the convergence of the ergodic means $M(f, \cdot)$ to a function $\mu : \Omega \to \mathbb{R}$ as in Definition 6.3 for state functions.

EXAMPLE. The ratios of the form $r_{t,t'}(v)$ defined in (27) can be expressed by ergodic means of extended state functions. Indeed, consider the extended state functions N_t and $N_{t'}$ defined by:

$$\forall \mathbf{s} \in \Sigma, \quad \forall z \in \Omega_{\mathbf{s}}, \quad N_t(\mathbf{s}, z) = \mathrm{Card}\{e \in z \,:\, \lambda(e) = t\}, \tag{28}$$

and similarly for $N_{t'}$. Then, for v finite and R-stopped, the ratio $r_{t,t'}(v)$ is given by:

$$r_{t,t'}(v) = \frac{\langle N_t, v \rangle}{\langle N_{t'}, v \rangle} = \frac{\langle N_t, v \rangle}{\langle \mathbf{1}, v \rangle} \times \frac{\langle \mathbf{1}, v \rangle}{\langle N_{t'}, v \rangle} = \frac{M(N_t, v)}{M(N_{t'}, v)}. \tag{29}$$

Hence, if we know that ergodic means of extended state functions have a limit, we shall conclude that the ratios $r_{t,t'}(v)$ also have a limit. This is the topic of the following result, which proof is postponed in §7.

Theorem 6.2 (extended LLN)   *Let $\big(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma}\big)$ be a Markov net, that we assume recurrent and with integrable concurrency height. Let $(\alpha(\mathbf{s}))_{\mathbf{s} \in \Sigma}$ denote the stationary measure of the Markov net.*

Then for every extended state function f, the ergodic means $M(f, \cdot)$ converge to a function $\mu(f, \cdot) : \Omega \to \mathbb{R}$. Moreover, $\mu(f, \cdot) = \mu(f)$ is constant with probability 1, given by:

$$\mu(f) = \sum_{\mathbf{s} \in \Sigma} \alpha(\mathbf{s}) p_{\mathbf{s}}(f), \qquad with: \quad p_{\mathbf{s}}(f) = \sum_{z \in \Omega_{\mathbf{s}}} p_{\mathbf{s}}(z) f(\mathbf{s}, z), \quad \mathbf{s} \in \Sigma. \tag{30}$$

Remark that, in case of an extended state function defined by a state function, the expression (30) giving $\mu(f)$ extends the expression (24) of the first LLN. Indeed, since $f(\mathbf{s}, \cdot) = f(\mathbf{s})$ is constant for every $\mathbf{s}$, and since $p_{\mathbf{s}}$ sums up to 1, $p_{\mathbf{s}}(f) = f(\mathbf{s})$ and thus $\mu(f) = \sum_{\mathbf{s} \in \Sigma} \alpha(\mathbf{s}) f(\mathbf{s})$.

We shall now complete the example of ratios $r_{t,t'}$ defined in (27). Using Equations (29) and (30), we have for every regular sequence $(V_n)_{n \geq 0}$ of stopping operators, and with probability 1:

$$\lim_{n \to \infty} r_{t,t'}(V_n) = \lim_{n \to \infty} \frac{M(N_t, V_n)}{M(N_{t'}, V_n)} = \frac{\mu(N_t)}{\mu(N_{t'})},$$

where N_t is defined in (28), and $N_{t'}$ is defined similarly. Hence, the ratios $R_{t,t'}$ have a limit with probability 1, and this limit is the same, with probability 1, for all $\omega \in \Omega$.

REMARK. It is not clear how to extend the LLN, for example, to functions defined on "pairs of successive clusters". This is done for a Markov chain $(X_n)_{n \geq 0}$ by considering $(X_n, X_{n-1})_{n \geq 1}$, which is again a Markov chain. Our attempts to get a similar construction for nets were not fruitful.

7   Proof of the Law of Large Numbers

The aim of this section is to prove the LLN, as stated in Theorems 6.1 and 6.2. The outline of the proof of Theorem 6.1 is as follows:

1. We first fix a particular sequence of stopping operators and examine the convergence of ergodic means along this particular sequence. We show that these ergodic means can be expressed by ergodic means associated with some homogeneous (usual) Markov chain, defined over some special, huge, state space. This shows the convergence of ergodic means to a function $\mu : \Omega \to \mathbb{R}$ for this particular sequence (Lemma 7.5).

2. We show that the limit obtained remains unchanged if we modify the regular sequence of stopping operators, and that the function μ is constant with probability 1 on Ω. This yields the density coefficients $(\alpha(\mathbf{s}))_{\mathbf{s} \in \Sigma}$ (Lemma 7.5).

3. We show that the coefficients are positive, which completes the proof of Theorem 6.1 (Lemma 7.6).

4. Theorem 6.2 is then obtained as a corollary of Theorem 6.1.

We stress the following facts: Theorem 6.1 does not use the properties of *distributed probabilities*: only the homogeneity stated in Theorem 5.1 is actually used. In contrast, the properties of distributed probabilities are used in the proof of Theorem 6.2.

Prior to proceeding with the above steps of the proof, we need to introduce some new notions, that are the topic of §§7.1–7.4. Then we focus in §§ 7.5 on the proofs of the theorems.

7.1   The Embedded Markov Chain

We first associate to each Markov net a finite Markov chain that "codes", in a sense to be made precise, the probabilistic behavior of the net. The chain is defined on a very large state space. It is thus of little interest in practice; but it has a theoretical merit, namely, in relating Markov nets to finite Markov chain theory.

Denote by Q the finite set consisting of all pairs (m, z), where $m \in \mathbf{M}_{\mathrm{rs}}$ and $z \in \Omega_{B_0(E^m)}$. That is, z is maximal in the max-initial prefix of E^m. Recall form §4.1 the definition of isomorphism of probability spaces.

Lemma 7.1   *Let $\big(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma}\big)$ be a Markov net, with Q the finite set defined as above, and let $(V_n)_{n \geq 0}$ and $(Z_n)_{n > 0}$ form the normal decomposition of maximal configurations. Then the sequence $(\gamma(V_n), Z_{n+1})_{n \geq 0}$, seen as a sequence of random variables with values in Q, defines a finite Markov chain.*

If $(\Xi, \mathfrak{G}, \mathbb{Q})$ denotes the canonical probability space associated with the Markov chain, there is a natural isomorphism of probability spaces $(\Omega, \mathfrak{F}, \mathbb{P}) \to (\Xi, \mathfrak{G}, \mathbb{Q})$, where $\mathfrak{F}$ denotes the Borel σ-algebra of Ω.

Proof. Recall the notations γ^m, $\mathbb{P}^m$, etc ... , used in reference with the net $\mathcal{N}^m$, for any reachable marking m. For any $m \in \mathbf{M}_{\mathrm{rs}}$, let also Z^m be the random variable $Z^m : \Omega_{E^m} \to \Omega_{B_0(E^m)}$, $\xi \in \Omega_{E^m} \mapsto Z^m(\xi) = \xi \cap B_0(E^m)$. We define a transition matrix P on Q as follows:

$$\forall (m, z), (m', z') \in Q, \quad P_{(m,z),(m',z')} = \mathbf{I}\big(\gamma^m(z) = m'\big) \mathbb{P}^{m'}(Z^{m'} = z'), \qquad (31)$$

where $\mathbf{I}(\text{PRED})$ takes the value 1 if the predicate PRED is true, 0 otherwise. It is clear that P is indeed a stochastic matrix on Q, i.e.:

$$\forall (m,z) \in Q, \quad \sum_{(m',z') \in Q} P_{(m,z),(m',z')} = 1.$$

For any $n \geq 1$, let $h_n(m_0, z_1, \ldots, m_{n-1}, z_n)$, with $(m_i, z_{i+1}) \in Q$ for all $i = 1, \ldots, n-1$, denote the following quantity:

$$h_n(m_0, z_1, \ldots, m_{n-1}, z_n) =$$
$$\mathbb{P}\big(\gamma(V_0) = m_0, Z_1 = z_1, \ldots, \gamma(V_{n-1}) = m_{n-1}, Z_n = z_n\big).$$

We have, by the chain rule:

$$\begin{aligned}
h_n(m_0, z_1, \ldots, m_{n-1}, z_n) &= h_{n-1}(m_0, z_1, \ldots, m_{n-2}, z_{n-1}) \\
&\quad \times \mathbb{P}^{\gamma^{m_{n-2}(z_{n-1})}}\big(m(V_1) = m_{n-1}, Z^{m_{n-1}} = z_n\big) \\
&= h_{n-1}(m_0, z_1, \ldots, m_{n-2}, z_{n-1}) \\
&\quad \times \mathbf{I}\big(\gamma^{m_{n-2}(z_{n-1})} = m_{n-1}\big) \times \mathbb{P}^{m_{n-1}}\big(Z^{m_{n-1}} = z_n\big) \\
&= h_{n-1}(m_1, z_1, \ldots, m_{n-1}, z_{n-1}) \times P_{(m_{n-2}, z_{n-1}),(m_{n-1}, z_n)}.
\end{aligned}$$

This shows that the sequence $(\gamma(V_n), Z_n)_{n>0}$ is a Markov chain with transition matrix P. The initial distribution μ of the chain is given by:

$$\forall (m,z) \in Q, \quad \mu(m,z) = \mathbf{I}\big(m = m_0\big)\mathbb{P}(Z_1 = z), \tag{32}$$

where m_0 is the initial marking of the net.

Let $(\Xi, \mathfrak{G}, \mathbb{Q})$ be the canonical sample space associated with the Markov chain. We have by construction a measurable map $\Phi : \Omega \to \Xi$, given by $\Phi(\omega) = \big(\gamma(V_n(\omega)), Z_{n+1}(\omega)\big)_{n \geq 0}$, such that $\Phi\mathbb{P} = \mathbb{Q}$, where $\Phi\mathbb{P}$ is the image probability of $\mathbb{P}$ by Φ (see § 4.1 for the definition of image probability). Let us show that Φ is injective. Indeed, if $\Phi(\omega) = \big(m_n, Z_{n+1}\big)_{n \geq 0}$, then the normal decomposition $(V_n, Z_n)_{n>0}$ of ω is entirely determined by $(Z_n)_{n>0}$, since $V_0 = \emptyset$, and $V_n = Z_1 \oplus \cdots \oplus Z_n$ for $n > 0$. Since $\omega = \sup_n V_n$ by Theorem 3.2, ω is entirely determined by $\Phi(\omega)$, which shows that Φ is injective. We finally show that Φ is onto. To this end, let $(m_n, z_{n+1})_{n \geq 0}$ be an element of Ξ. Consider the following sequence:

$$v_0 = \emptyset, \qquad n > 0, \quad v_n = v_{n-1} \oplus z_n.$$

We show by induction on n that v_n is a configuration of E, with $\gamma(v_n) = m_n$. This is trivial for $n = 0$. To see it for $n = 1$, we may assume without loss of generality that $\mu(m_0, z_1) > 0$, since:

$$\mathbb{Q}\big(\mu(m_0, z_1) = 0\big) = 0.$$

According to (32), this implies in particular that $\mathbf{I}(m = m_0) = 1$, so that $z_1 \in \Omega_{B_0(E)}$. Hence $v_1 = z_1$ is a configuration of E. For the same reason, we may assume that $P_{(m_0,z_1),(m_1,z_2)} > 0$. According to (31), this implies that $\mathbf{I}\big(\gamma^{m_0}(z_1) = m_1\big) = 1$, and thus $\gamma(v_1) = m_1$, which shows the induction hypothesis for $n = 1$. The general case follows along the same line. Hence $(v_n)_{n \geq 0}$ is a nondecreasing sequence of configurations of E. Therefore $v = \sup_{n \geq 0} v_n$ is also a configuration of E. Let ω be any maximal configuration of E containing v. Then it is clear, by induction on $n \geq 0$, that the normal decomposition of ω satisfies $V_n(\omega) = v_n$ for $n \geq 0$, $Z_n(\omega) = z_n$ for $n > 0$. Hence, by Theorem 3.2, $\omega = \sup_n v_n = v$. Therefore, ω satisfies $\Phi(\omega) = (m_n, z_{n+1})_{n \geq 0}$. This shows hat Φ is onto a set of $\mathbb{Q}$-probability 1, and completes the proof (since Φ is clearly bi-measurable). $\square$

COMMENT. Lemma 7.1 is surprising: finally, a Markov net is nothing but a special Markov chain, defined on a—huge—finite set. However, the concurrency properties of the net are hidden in the Markov chain representation, whereas they are clearly revealed by the net representation. Our work indeed aims at revealing the concurrency properties of the model from the probabilistic viewpoint.

Note that this special Markov chain is still different from the usual one associated with the Marking graph of the net. Our former Markov chain makes the true-concurrent probabilistic semantics "rigid", whereas the latter relates to the interleaving semantics.

Remark also that the lemma only uses the homogeneity property (16); the fact that $\mathbb{P}$ is a distributed product is not essential here, it is only a sufficient condition for the homogeneity. Hence the same result holds even without the local finiteness assumption, provided that the probability has the homogeneity property. In turn, the associated Markov chain would be, in general, defined on an *infinite* state space. This enlightens the role of the local finiteness assumption.

7.2 Preliminaries on Global Recurrence

Recall that we have defined in §5.3 the return operator R associated to a Markov net. Also, the successive returns are defined by:

$$R_0 = \emptyset, \qquad \forall n > 0, \quad R_n(\omega) = R_{n-1}(\omega) \oplus R\big(\omega \ominus R_{n-1}(\omega)\big). \tag{33}$$

In case of a Markov net that reduces to a Markov chain, R_n coincides with the nth return to the initial state. These are known to be stopping times. For general Markov nets, we have:

Lemma 7.2 *If $\mathcal{N}$ is recurrent, the successive return operators R_n form a regular sequence of stopping operators.*

Proof. The fact that every R_n is a stopping operator follows by induction from the fact that R is a stopping operator. But this follows easily from the definitions (see a detailed proof in [1, Ch.6, I-1.*6*] or in [2]).

Let ω be an element of Ω such that all $R_n(\omega)$ are well-defined, $n \geq 0$. We have $\bigcup_{n \geq 0} R_n(\omega) \subseteq \omega$ by construction; for the converse inclusion, let $v = \bigcup_{n \geq 0} R_n(\omega)$, and assume that v is not maximal. Then there is an event e minimal in E^v. But then there is an integer n such that e is minimal in $E^{R_p(\omega)}$ for all $p \geq n$. In particular, e is a minimal event of both $E^{R_n(\omega)}$ and $E^{R_{n+1}(\omega)}$, which contradicts the definition of R. Hence v is maximal and thus $v = \omega$. Finally it is clear that $\langle \mathbf{1}, R \rangle \geq 1$, and thus by induction, using (33), $\langle \mathbf{1}, R_n \rangle \geq n$. This shows that $(R_n)_{n \geq 0}$ is a regular sequence of stopping operators, and completes the proof of the lemma. $\qquad\square$

Recall that if $\varphi : P \to Q$ is a measurable map, where P and Q are two sets respectively equipped with the σ-algebras $\mathfrak{F}$ and $\mathfrak{G}$, the σ-algebra $\langle \varphi \rangle$ generated by φ is defined by $\langle \varphi \rangle = \{\varphi^{-1}(A), A \in \mathfrak{G}\}$, and then $\langle \varphi \rangle \subseteq \mathfrak{F}$.

Let $\mathcal{N}$ be a recurrent net. We may assume without loss of generality that $R_n(\omega)$ is finite for every $n \geq 0$ (e.g., by redefining $R(\omega) = \emptyset$ whenever $R(\omega) = \omega$, note that the set of all these ω has probability 0). Hence R_n takes its values in a set at most countable. Therefore the associated σ-algebra $\langle R_n \rangle$ can be described as follows:

$$\langle R_n \rangle = \langle R_n^{-1}(u), \, u \in \mathcal{W}_E \rangle \, ;$$

that is, $\langle R_n \rangle$ is the smallest σ-algebra that contains the family of subsets $\{R_n^{-1}(u), \, u \in \mathcal{W}_E\}$.

Furthermore, we note the following property of R_n's, which is a general property of stopping operators: If u is a finite R-stopped configuration such that $u = R_n(\omega_o)$ for some $n \geq 0$ and $\omega_o \in \Omega$, then we have:

$$R_n^{-1}(u) = \{\omega \in \Omega \, : \, \omega \supseteq u\}. \tag{34}$$

Indeed, since R_n is a stopping operator by Lemma 7.2, it is enough to verify (34) for stopping operators; but this is an immediate consequence of point 2 in Definition 6.2.

Corollary 7.1   *Assume that $\mathcal{N}$ is recurrent, and denote by $\mathfrak{F}$ the Borel σ-algebra on Ω, and for all $n \geq 0$, by $\mathfrak{F}_n$ the σ-algebra generated by R_n. Then $\mathfrak{F} = \langle \mathfrak{F}_n, \, n \geq 0 \rangle$, i.e., $\mathfrak{F}$ coincides with the smallest σ-algebra that contains all $\mathfrak{F}_n$, $n \geq 0$.*

Proof. Let $\mathfrak{G} = \langle \mathfrak{F}_n, \, n \geq 0 \rangle$. We obviously have $\mathfrak{G} \subseteq \mathfrak{F}$. For the converse inclusion, recall the notation $\mathcal{S}(u) = \{\omega \in \Omega \, : \, \omega \supseteq u\}$. Since $\mathfrak{F}$ is generated by the collection $\{\mathcal{S}(u)\}$, where u ranges over the set $\mathcal{V}_E$ of finite configurations of E, it is enough to show that $\mathcal{S}(u) \in \mathfrak{G}$ for every $u \in \mathcal{V}_E$.

Let $K = \{R_n(\omega), \, n \geq 0, \, \omega \in \Omega\}$. K is at most countable, since it consists of finite configurations. Let $u \in \mathcal{V}_E$. From the equality $\omega = \sup_{n \geq 0} R_n(\omega)$ stated in Lemma 7.2, we get:

$$\mathcal{S}(u) = \bigcup_{\substack{v \in K \\ v \supseteq u}} \mathcal{S}(v) \, .$$

Using (34), each $\mathcal{S}(v)$ with $v \in K$ can be written $\mathcal{S}(v) = R_{n_v}^{-1}(v)$, where $n_v \geq 0$. Thus:

$$\mathcal{S}(u) = \bigcup_{\substack{v \in K \\ v \supseteq u}} R_{n_v}^{-1}(v) \,,$$

a union at most countable of $\mathfrak{G}$-measurable subsets. Hence $\mathcal{S}(u) \in \mathfrak{G}$, which completes the proof of the corollary. $\square$

7.3 Preliminaries on Local Recurrence

Fix $\mathbf{s}_0$ a dynamic cluster of $\mathcal{N}$, and consider $\omega \in \Omega$. Assume that x and x' are two branching cells in $\Delta(\omega)$ such that $\langle x \rangle = \langle x' \rangle = \mathbf{s}_0$. Then there are events in x and x' that are both compatible, and labeled by the same transition. Since the net is safe, this implies that these two events are causally related. This induces in turn an ordering of such branching cells, for $\mathbf{s}_0$ and ω fixed, which is thus a total ordering. Furthermore, for each branching cell $x \in \Delta(\omega)$, the configuration:

$$v_x = \inf\{v \in \mathcal{W}, \, v \subseteq \omega, \, x \in \delta_E(v)\}, \tag{35}$$

is a finite R-stopped configuration, since this set of configurations is nonempty, and by the lattice property of compatible R-stopped sub-configurations of ω (Theorem 3.1, point 5). The above ordering on branching cells x such that $x \in \Delta(\omega)$ and $\langle x \rangle = \mathbf{s}_0$ corresponds to the set-inclusion of associated configurations v_x.

In particular, again with ω fixed, if the following set is nonempty:

$$\{x \in \Delta(\omega) \, : \, \langle x \rangle = \mathbf{s}_0, \, x \notin \delta_E(\emptyset)\},$$

it has a unique minimal element x, with an associated configuration v_x defined as in (35). Let $S^{\mathbf{s}_0}(\omega)$ denote this configuration, so that $S^{\mathbf{s}_0}(\omega)$ is defined by:

$$S^{\mathbf{s}_0}(\omega) = \inf\{v \in \mathcal{W} \, : \, v \subseteq \omega, \, v \neq \emptyset, \, \exists x \in \delta_E(v), \, \langle x \rangle = \mathbf{s}_0\}. \tag{36}$$

Remark that the branching cell x in (36) is then unique, again for safeness reasons. Moreover, since the net is assumed recurrent, a simple Borel-Cantelli argument shows that $S^{\mathbf{s}_0}$ is defined on Ω with probability 1.

Definition 7.1 (local return operator) *If the Markov net $(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma})$ is recurrent, for every dynamic cluster $\mathbf{s}_0$, the* local return to $\mathbf{s}_0$ *is the mapping $S^{\mathbf{s}_0} : \Omega \rightarrow \mathcal{W}$ defined by (36), with probability 1.*

Intuitively, local returns are returns with "minimal moves of the tokens". For every reachable marking m, and in particular if m is recursively stopped, the same definition of local return operator applies to Ω_{E^m}. Since $\mathcal{N}^m$ is itself also recurrent, the local return to $\mathbf{s}_0$ is also defined on Ω_{E^m} with $\mathbb{P}^m$-probability 1. We denote this local return operator by:

$$S^{\mathbf{s}_0, m} : \Omega_{E^m} \rightarrow \mathcal{W}_{E^m}, \quad \text{with } \mathbb{P}^m\text{-probability } 1. \tag{37}$$

This allows to construct the *successive local return operators to* $\mathbf{s}_0$ as follows. The local returns $(S_n^{\mathbf{s}_0})_{n \geq 1}$ to $\mathbf{s}_0$ are defined on Ω with probability 1 by:

$$S_1^{\mathbf{s}_0} = S^{\mathbf{s}_0}, \qquad \forall n > 1, \qquad S_n^{\mathbf{s}_0}(\omega) = S_{n-1}^{\mathbf{s}_0}(\omega) \oplus S^{\mathbf{s}_0, m_{n-1}}\big(\omega \ominus S_{n-1}^{\mathbf{s}_0}(\omega)\big),$$
$$\text{with} \quad m_{n-1} = \gamma\big(S_{n-1}^{\mathbf{s}_0}(\omega)\big). \tag{38}$$

We summarize the properties of local return operators as follows:

Proposition 7.1   *Let* $\big(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma}\big)$ *be a recurrent Markov net, and let* $\mathbf{s}_0$ *be a dynamic cluster of* $\mathcal{N}$. *Then for each* $n \geq 1$, *the* n^{th} *local return operator* $S_n^{\mathbf{s}_0}$ *to* $\mathbf{s}_0$ *is a stopping operator of* Ω. *If* $\big(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s} \in \Sigma}\big)$ *has integrable concurrency height, then* $(S_n^{\mathbf{s}_0})_{n \geq 1}$ *is a regular sequence of stopping operators.*

Proof. As for global return operators, it is enough to show that the single $S^{\mathbf{s}_0}$ defined by (36) is a stopping operator. To this end, we clearly have $S^{s_0}(\omega) \subseteq \omega$. Let $\omega, \omega' \in \Omega$, set $v = S^{s_0}(\omega)$, $v' = S^{s_0}(\omega')$, and assume that $\omega' \supseteq v$. Let x be the unique element of $\delta_E(v)$ such that $\langle x \rangle = \mathbf{s}_0$. Then, since v is finite nonempty R-stopped, since $x \in \delta_E(v)$ and since $v \subseteq \omega'$, we have $v' \subseteq v$ by the very definition of $v' = S^{\mathbf{s}_0}(\omega')$. This implies $v' \subseteq \omega$, and in turn, by minimality of $S^{\mathbf{s}_0}(\omega)$, we get $v \subseteq v'$. Finally, $v = v'$, which proves that $S^{\mathbf{s}_0}$ is a stopping operator.

Assume moreover that the net has integrable concurrency height; we show that the sequence $(S_n^{\mathbf{s}_0})_{n \geq 1}$ of local returns to $\mathbf{s}_0$ is regular. $(S_n^{\mathbf{s}_0})_{n \geq 1}$ is clearly nondecreasing. Since $S^{\mathbf{s}_0} \neq \emptyset$, $\langle N, S(\omega) \rangle \geq 1$ with probability 1; it follows thus from (38) that $\langle N, S_n^{\mathbf{s}_0} \rangle \geq n$ for all $n \geq 1$. Hence it remains only to show that $\sup_{n \geq 1} S_n^{\mathbf{s}_0}(\omega) = \omega$ with probability 1. Fix $\omega \in \Omega$, let $v_n = S_n^{\mathbf{s}_0}(\omega)$, $v = \sup_{n \geq 1} v_n$, and assume that $v \neq \omega$. Consider any minimal event $e \in E^v$. Then there is an integer p such that e is also a minimal event of E^{v_n} for all $n \geq p$. Pick any place b in the preset of $\lambda(e)$, where $\lambda(e)$ is the transition that labels e. For $m = \gamma(v_p)$, the token in the place b is a frozen token, from v_p to ω, i.e., the length $\langle N, \omega \ominus v_p \rangle$ is infinite. But, since the length is integrable, all such ω have together probability 0. This shows that $\sup_{n \geq 1} S_n^{\mathbf{s}_0}(\omega) = \omega$ with probability 1, and completes the proof of the proposition. □

Still consider a fixed dynamic cluster $\mathbf{s}_0$, and let $\omega \in \Omega$ be such that the local returns $S_n^{\mathbf{s}_0}(\omega)$ to $\mathbf{s}_0$ are defined for all $n \geq 1$, which holds with probability 1 if the net is recurrent. For each $n \geq 1$, there is by construction a unique branching cell $x_n(\omega) \in \delta_E\big(S_n^{\mathbf{s}_0}(\omega)\big)$ such that $\langle x_n(\omega) \rangle = \mathbf{s}_0$. As a consequence of Lemma 3.1, we have $\omega \cap x_n(\omega) \in \Omega_x = \Omega_{\mathbf{s}_0}$. Therefore, if we set:

$$\forall n \geq 1, \quad Y_n^{\mathbf{s}_0}(\omega) = \omega \cap x_n(\omega), \tag{39}$$

we get a sequence $(Y_n^{\mathbf{s}_0})_{n \geq 1}$ of random variables such that:

$$\forall n \geq 1, \quad Y_n^{\mathbf{s}_0}(\omega) \in \Omega_{\mathbf{s}_0}, \quad \text{with probability 1.}$$

Using 1) the Strong Markov property for Markov nets, and 2) the properties of the distributed product, yields the following result regarding the sequence $(Y_n^{\mathbf{s}_0})_{n \geq 1}$ (see [1, Ch.6, III-3.2] for the proof):

Lemma 7.3 *Let $\big(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s}\in\Sigma}\big)$ be a recurrent Markov net, let $\mathbf{s}_0$ be a dynamic cluster, and let $(Y_n^{\mathbf{s}_0})_{n\geq 1}$ be the sequence of random variables, with values in $\Omega_{\mathbf{s}_0}$, defined by (39). Then $(Y_n^{\mathbf{s}_0})_{n\geq 1}$ is a sequence of independent and identically distributed variables, with law $p_{\mathbf{s}_0}$ in $\Omega_{\mathbf{s}_0}$.*

We finally show the following result, which generalizes to Markov nets the *positive recurrence* of recurrent states, for Markov chains:

Lemma 7.4 (positive recurrence) *Let $\big(\mathcal{N}, (p_{\mathbf{s}})_{\mathbf{s}\in\Sigma}\big)$ be a recurrent Markov net. Then the global return R, and all the local returns $S^{\mathbf{s},m}$, $\mathbf{s}\in\Sigma$, $m\in\mathbf{M}_{\mathrm{rs}}$, satisfy:*

$$\mathbb{E}\big(\langle \mathbf{1}, R\rangle\big) < \infty, \qquad \mathbb{E}^m\big(\langle \mathbf{1}, S^{\mathbf{s},m}\rangle\big) < \infty,$$

where $\mathbb{E}$ and $\mathbb{E}^m$ denote respectively the mathematical expectation w.r.t. probabilities $\mathbb{P}$ and $\mathbb{P}^m$.

Proof. We first show that $\mathbb{E}(\langle \mathbf{1}, R\rangle) < \infty$. Let T be the integer random variable $T = \langle \mathbf{1}, R\rangle$. Recall the usual identity $\mathbb{E}(T) = \sum_{n\geq 0}\mathbb{P}(T \geq n)$. Assume that there is an integer $r > 0$ such that:

$$\sum_{n\geq 0}\mathbb{P}(T \geq nr) < \infty. \tag{40}$$

Then we have:

$$\mathbb{E}(T) = \sum_{n\geq 0}\mathbb{P}(T \geq n) = \sum_{i=0}^{\infty}\sum_{j=0}^{r-1}\mathbb{P}(T \geq ir + j)$$
$$\leq \sum_{i=0}^{\infty}\sum_{j=0}^{r-1}\mathbb{P}(T \geq ir)$$
$$= r\sum_{n=0}^{\infty}\mathbb{P}(T \geq nr) < \infty, \qquad \text{by (40)}.$$

Hence, to prove $\mathbb{E}(T) < \infty$, it is enough to show (40). Let $(V_n)_{n\geq 0}$ be the sequence of stopping operators coming from the normal decomposition of maximal configurations, as defined in §3.6 and in §7.1, Lemma 7.1. There is an integer $k_1 > 0$ such that $\langle \mathbf{1}, V_n\rangle \leq k_1 \times n$ for all $n > 0$; take for example k_1 as the maximal number of simultaneously enabled transitions of the net. Define the random variable T' by:

$$T' = \inf\{n \geq 0 \,:\, R \subseteq V_n\}.$$

Since R is finite with probability 1, and since $\bigcup_n V_n(\omega) = \omega$ for all $\omega \in \Omega$, T' is finite with probability 1. Therefore, with probability 1, $T \leq \langle \mathbf{1}, V_{T'}\rangle \leq k_1 \times T'$. Hence, to show (40), it is enough to show that there is an integer $q > 0$ such that:

$$\sum_{n\geq 0}\mathbb{P}(T' \geq qn) < \infty. \tag{41}$$

We choose the integer q as follows. Recall that $\mathbf{M}_{\mathrm{rs}}$ denotes the set of recursively stopped markings of the net. Since the net is recurrent, we choose for any marking $m \in \mathbf{M}_{\mathrm{rs}}$ a finite R-stopped configuration v_m leading back from m to the initial marking, after having moved all tokens in the net. Each v_m has $\mathbb{P}^m$-positive likelihood. We put $q = \max_{m \in \mathbf{M}_{\mathrm{rs}}} \langle \mathbf{1}, v_m \rangle$, which is finite since $\mathbf{M}_{\mathrm{rs}}$ is a finite set, and positive otherwise we would have $v_m = \emptyset$ for all $m \in \mathbf{M}_{\mathrm{rs}}$. Let n be any integer $n \geq 1$, and let $\mathbb{Q}$ be the conditional probability $\mathbb{Q} = \mathbb{P}\big(\cdot \, | T' \geq q(n-1) \big)$. Then we have, by the Bayes rule:

$$\mathbb{Q}(T' \geq qn) = \sum_{m \in \mathbf{M}_{\mathrm{rs}}} \mathbb{Q}\big(T' \geq qn \, | \, \gamma(V_{q(n-1)}) = m \big) \times \mathbb{Q}\big(\gamma(V_{q(n-1)} = m \big), \qquad (42)$$

where $\gamma(v)$ denotes the marking reached a configuration v. Denote, for any $m \in \mathbf{M}_{\mathrm{rs}}$, by $(V_j^m)_{j \geq 0}$ the sequence of stopping operators coming from the normal decomposition of maximal configurations, defined on Ω_{E^m}. Then, using the homogeneity property (Theorem 5.1), $\mathbb{Q}\big(T' \geq qn \, | \, \gamma(V_{q(n-1)}) = m \big)$ is the $\mathbb{P}^m$ probability that V_q^m does not contain any return to the initial marking. By definition of q, this probability is less than a constant $a < 1$. Hence we get from (42) that $\mathbb{Q}(T' \geq qn) \leq a$, and thus, coming back to the definition of $\mathbb{Q}$:

$$\mathbb{P}(T' \geq qn) \leq a\mathbb{P}\big(T' \geq q(n-1)\big) \leq a^2 \mathbb{P}\big(T' \geq q(n-2)\big) \leq \cdots \leq a^n.$$

Since $a < 1$, Equation (41) follows, hence $\mathbb{E}(T) < \infty$.

We now show that $\mathbb{E}(\langle \mathbf{1}, S^{\mathbf{s},m} \rangle) < \infty$ for every $\mathbf{s} \in \Sigma$ and $m \in \mathbf{M}_{\mathrm{rs}}$. Since the net $\mathcal{N}^m$ is recurrent by Lemma 5.2, we may assume without loss of generality that $m = m_0$ is the initial marking. If there is a branching cell x such that: x is an initial stopping prefix of E, and such that $\langle x \rangle = \mathbf{s}$, then $S^{\mathbf{s},m_0} \subseteq R$. Therefore $\langle \mathbf{1}, S^{\mathbf{s},m_0} \rangle \leq \langle \mathbf{1}, R \rangle$ and thus $\mathbb{E}(\langle \mathbf{1}, S^{\mathbf{s},m_0} \rangle) \leq \mathbb{E}(\langle \mathbf{1}, R \rangle) < \infty$.

Finally we show that the general case reduces to this particular case. For any $M \in \mathbf{M}_{\mathrm{rs}}$, define the M-reachability operator A^M by:

$$\forall \xi \in \Omega_{E^m}, \quad A^M(\omega) = \min\{v \in \mathcal{W}_E \, : \, v \subseteq \xi, \, \gamma(v) = M, \, \mathrm{Min}(E) \cap \mathrm{Min}(E^v) = \emptyset\}.$$

Then A^M is finite with probability 1, and using the very same technique than above, we conclude that $\mathbb{E}(\langle \mathbf{1}, A^M \rangle) < \infty$. Now for any $\mathbf{s} \in \Sigma$, let $M \in \mathbf{M}_{\mathrm{rs}}$ such that there is a branching cell x and a $v \in \mathcal{W}_E$ with $\gamma(v) = M$, $\langle x \rangle = \mathbf{s}$ and $x \in \delta_E(v)$. Then we have:

$$\forall \omega \in \Omega, \quad S^{\mathbf{s},m_0}(\omega) \subseteq A^M(\omega) \oplus S^{M,\mathbf{s}}\big(\omega \ominus A^M(\omega)\big).$$

Therefore:

$$\mathbb{E}(\langle \mathbf{1}, S^{\mathbf{s},m_0} \rangle) \leq \mathbb{E}(\langle \mathbf{1}, A^M \rangle) + \mathbb{E}^M\big(\langle \mathbf{1}, S^{\mathbf{s},M} \rangle\big).$$

We have $\mathbb{E}(\langle \mathbf{1}, A^M \rangle) < \infty$ as we remarked above, and $\mathbb{E}^M\big(\langle \mathbf{1}, S^{\mathbf{s},M} \rangle\big) < \infty$ by the first case. Therefore $\mathbb{E}(\langle \mathbf{1}, S^{\mathbf{s},m_0} \rangle) < \infty$, which completes the proof of the lemma. $\qquad \square$

7.4 Preliminaries on Ergodicity

Definition 7.2 (homogeneous function) *Let* $\mathbf{M}_{\mathrm{rs}}$ *denote the set of recursively stopped markings of* $\mathcal{N}$ —*i.e., those markings reached by some finite R-stopped configurations. A family* $H = (H^m)_{m \in \mathbf{M}_{\mathrm{rs}}}$ *of real-valued functions* $H^m : \Omega_{E^m} \to \mathbb{R} \cup \{\infty\}$ *is called a* homogeneous function.

The homogeneous function H *is said to be nonnegative, respectively integrable, if every* H^m *is nonnegative, respectively integrable w.r.t. probability* $\mathbb{P}^m$.

If $H = (H^m)_{m \in \mathbf{M}_{\mathrm{rs}}}$ *is a homogeneous function, we consider the family* $(H^v)_{v \in \mathcal{W}_E}$ *of real-valued functions* $H^v : \Omega_{E^v} \to \mathbb{R} \cup \{\infty\}$ *defined by:*

$$\forall v \in \mathcal{W}_E, \quad H^v = H^{\gamma(v)},$$

and that satisfies:

$$\forall v, v' \in \mathcal{W}, \quad \gamma(v) = \gamma(v') \Rightarrow H^v = H^{v'}.$$

The two representations $(H^m)_{m \in \mathbf{M}_{\mathrm{rs}}}$ *and* $(H^v)_{v \in \mathcal{W}}$ *are obviously equivalent.*

We now prove the following result, to be interpreted as an *ergodicity result*.

Theorem 7.1 *Let* $\mathcal{N}$ *be a recurrent Markov net. Let* $H = (H^v)_{v \in \mathcal{W}}$ *be a nonnegative and integrable homogeneous function. Assume that we have, with probability* 1:

$$\forall \omega \in \Omega, \quad \forall v \in \mathcal{W}_E, \quad v \subseteq \omega \Rightarrow H^v(\omega \ominus v) = H^{\emptyset}(\omega). \tag{43}$$

Then for every $v \in \mathcal{W}_E$, H^v *is constant on a set of probability* 1*, and all the so obtained constants are identical.*

Proof. We begin with the following observation: If V is a stopping operator, with $V(\omega)$ finite with probability 1, and if $\mathbb{E}(\,\cdot\,|\,V)$ denotes the conditional expectation w.r.t. the σ-algebra $\langle V \rangle$ generated by V (see §4.1), we have for every integrable function $f : \Omega \to \mathbb{R}$:

$$\mathbb{E}(f \mid V) = \int_{\Omega^{\gamma(V)}} f(V \oplus \xi) \, d\mathbb{P}^{\gamma(V)}(\xi), \tag{44}$$

where $V \oplus \xi$ denotes the concatenation of V and ξ (proof left to the reader; hint: use (34) with stopping operator V).

We show first that $H^{\emptyset}$ is constant on Ω with probability 1. By Corollary 7.1, we have $\mathfrak{F} = \langle \mathfrak{F}_n, n \geq 0 \rangle$, where $\mathfrak{F}_n = \langle R_n \rangle$. The Martingale convergence theorem [7, Th. 35.5] implies thus:

$$H^{\emptyset} = \lim_{n \to \infty} \mathbb{E}(H^{\emptyset} \mid \mathfrak{F}_n), \quad \text{with probability 1.} \tag{45}$$

We apply (44) with $f = H^{\emptyset}$ to get:

$$\begin{aligned} \mathbb{E}(H^{\emptyset} \mid \mathfrak{F}_n) &= \int_{\Omega^{\gamma(R_n)}} H^{\emptyset}(R_n \oplus \xi) \, d\mathbb{P}^{\gamma(R_n)}(\xi) \\ &= \int_{\Omega} H^{\emptyset}(R_n \oplus \xi) \, d\mathbb{P}(\xi), \end{aligned} \tag{46}$$

the latter by applying the homogeneity of $\mathbb{P}$ (Theorem 5.1), and since $\gamma(R_n) = m_0$ by construction of the successive return operators R_n. Applying the assumption (43) with $v = R_n$, we get:

$$\forall \xi \in \Omega, \quad H^\emptyset(R_n \oplus \xi) = H^{R_n}(\xi) = H^\emptyset(\xi), \tag{47}$$

the latter since H is a homogeneous function. Using together (45), (46) and (47), we obtain:

$$H^\emptyset(\omega) = \lim_{n \to \infty} \int_\Omega H^\emptyset(R_n \oplus \xi) \, d\mathbb{P}(\xi) = \int_\Omega H^\emptyset(\xi) \, d\mathbb{P}(\xi) = \mathbb{E}(H^\emptyset),$$

with probability 1. This shows that $H^\emptyset$ is constant on Ω with probability 1. For the same reasons, every H^v, with $v \in \mathcal{W}$, is constant on Ω^v with $\mathbb{P}^v$ probability 1. It follows from (43), and from the fact that the likelihood of every $v \in \mathcal{W}$ is positive, that all the constants coincide. $\qquad\square$

Finally, the following result on homogeneous functions will be useful. It is the adaptation of a classical lemma from dynamical systems theory; our proof is borrowed from [12].

Proposition 7.2 *Let $\mathcal{N}$ be a Markov net with unfolding E, and let $(V_n)_{n \geq 0}$ be a regular sequence of stopping operators on Ω. For each $n \geq 0$, we define θ_n by:*

$$\forall \omega \in \Omega, \quad \theta_n(\omega) = \omega \ominus V_n(\omega), \qquad so \ that: \ \theta_n \in \Omega_{E^{\gamma(V_n)}} \,.$$

Let also $H = (H^v)_{v \in \mathcal{W}}$ be a homogeneous, nonnegative and integrable function. Then we have:

$$\lim_{n \to \infty} \frac{H^{V_n}(\theta_n)}{\langle N, V_n \rangle} = 0, \quad with \ probability \ 1.$$

Proof. We also use the notation $(H^m)_{m \in \mathbf{M}_{\mathrm{rs}}}$ for H, where $\mathbf{M}_{\mathrm{rs}}$ is the set of recursively stopped markings. Since $(V_n)_{n \geq 0}$ is a regular sequence of operators, there is a constant $k > 0$ such that $\langle N, V_n \rangle \geq k \times n$ for every $n \geq 1$. Therefore it is enough to show:

$$\lim_{n \to \infty} \frac{H^{V_n}(\theta_n)}{n} = 0, \quad with \ probability \ 1.$$

To this end, we denote by X_n the random variable $X_n = \frac{1}{n} H^{V^n}(\theta_n)$, and we use the following classical criterion that implies the convergence of $(X_n)_{n \geq 1}$ to 0 with probability 1:

$$\forall \epsilon > 0, \quad \sum_{n \geq 1} \mathbb{P}(X_n \geq \epsilon) < \infty. \tag{48}$$

For each $n \geq 1$, let K_n denote the set of values of V_n. Since V_n is finite with probability 1, we assume without loss of generality that K_n is at most countable. Since V_n is a stopping operator, we have the property already observed:

$$\forall u \in K_n, \quad \{\omega \in \Omega \,:\, V_n(\omega) = u\} = \mathcal{S}(u),$$

where $\mathcal{S}(u)$ denotes as usual $\mathcal{S}(u) = \{\omega \in \Omega : \omega \supseteq u\}$. From this we get:

$$
\begin{aligned}
\mathbb{P}(X_n \geq \epsilon) &= \sum_{u \in K_n} \mathbb{P}(V_n = u)\mathbb{P}\big(H^{V_n}(\theta_n) \geq n\epsilon \,\big|\, V_n = u\big) \\
&= \sum_{u \in K_n} \mathbb{P}(V_n = u)\mathbb{P}^u\big(H^u \geq n\epsilon\big) \\
&= \sum_{m \in \mathbf{M}_{\mathrm{rs}}} \bigg(\sum_{\substack{u \in K_n \\ \gamma(u)=m}} \mathbb{P}(V_n = u) \bigg)\mathbb{P}^m\big(H^m \geq n\epsilon\big) \\
&\leq \sum_{m \in \mathbf{M}_{\mathrm{rs}}} \mathbb{P}^m\big(H^m \geq n\epsilon\big).
\end{aligned}
\tag{49}
$$

In order to show (48), and from (49), it is enough to show that for each $m \in \mathbf{M}_{\mathrm{rs}}$, the following sum is finite: $\sum_{n \geq 1} \mathbb{P}^m(H^m \geq n\epsilon) < \infty$. Recall the usual equality $\mathbb{E}(f) = \sum_{k=1}^{\infty} \mathbb{P}(f \geq k)$ for every nonnegative integrable function $f : \Omega \to \mathbb{N} \cup \{\infty\}$. With this transformation, we get for every $m \in \mathbf{M}_{\mathrm{rs}}$:

$$
\begin{aligned}
\sum_{n \geq 1} \mathbb{P}^m(H^m \geq n\epsilon) &= \sum_{n \geq 1} \mathbb{P}^m\big(\tfrac{1}{\epsilon}H^m \geq n\big) \\
&= \frac{1}{\epsilon}\mathbb{E}^m(H^m) < \infty,
\end{aligned}
$$

since H^m is integrable for every $m \in \mathbf{M}_{\mathrm{rs}}$. This completes the proof. $\qquad\square$

7.5 Proof of Theorems 6.1 and 6.2.

We begin with the proof of Theorem 6.1. We decompose the proof in two steps: first, the existence of the density coefficients (points 1 and 2 of the theorem, shown in Lemma 7.5 below), then their properties (point 3 of the theorem, shown in Lemma 7.6 below).

7.5.1 Existence of the Density Coefficients

The ergodic means $M(f, v)$ are linear in f. State functions form a vector space of finite dimension, with basis the collection of state function $\mathbf{1}_{\mathbf{s}_0}$, $\mathbf{s}_0 \in \Sigma$, defined by:

$$
\forall \mathbf{s} \in \Sigma, \quad \mathbf{1}_{\mathbf{s}_0}(\mathbf{s}) = \begin{cases} 1, & \text{if } \mathbf{s} = \mathbf{s}_0, \\ 0, & \text{otherwise.} \end{cases}
\tag{50}
$$

Therefore, to obtain the convergence of ergodic means of any state function f stated in (24) in Theorem 6.1, it is enough to show the following lemma:

Lemma 7.5 *For every dynamic cluster $\mathbf{s}_0$, the ergodic means $M(\mathbf{1}_{\mathbf{s}_0}, \cdot)$ converge in the sense of Definition 6.3 to a function $\alpha(\mathbf{s}_0, \cdot) : \Omega \to \mathbb{R}$. The function $\alpha(\mathbf{s}_0, \cdot)$ is constant with probability 1 on Ω.*

Proof. We fix a dynamic cluster $\mathbf{s}_0$ of $\mathcal{N}$, and we consider the state function $f_0 = \mathbf{1}_{\mathbf{s}_0}$ defined by (50).

Let $(M_n, Z_{n+1})_{n \geq 0}$ denote the embedded Markov chain of the net (see §7.1), where $M_n = \gamma(V_n)$ is the marking associated with the configuration V_n, arising from the normal decomposition of maximal configurations. The ergodic means relative to f_0 and V_n satisfy:

$$
\begin{aligned}
M(f_0, V_n) &= \frac{\langle f_0, V_n \rangle}{n} \frac{n}{\langle \mathbf{1}, V_n \rangle} \\
&= \underbrace{\frac{\sum_{k=0}^{n-1} \mathbf{I}\big(\mathbf{s}_0 \in \delta_{E^{M_k}}(\emptyset)\big)}{n}}_{\text{ergodic means for } (M_n, Z_n)} \cdot \underbrace{\left(\frac{\sum_{k=0}^{n-1} \mathrm{Card}\big(\delta_{E^{M_k}}(\emptyset)\big)}{n} \right)^{-1}}_{\text{ergodic means for } (M_n, Z_n)},
\end{aligned}
$$

where we recall that $\mathbf{I}(\text{PRED})$ takes the value 1 if the predicate PRED is true, 0 otherwise. Hence, each factor is given by ergodic means relative to the embedded Markov chain $(M_n, Z_{n+1})_{n \geq 0}$. The ergodic theory of Markov chains implies that each of these factors has a limit with probability 1 (see Theorem A.1 in Appendix A). Therefore, the following real-valued random variable $G : \Omega \to \mathbb{R}$ is well defined with probability 1:

$$
G(\omega) = \lim_{n \to \infty} M(f_0, V_n(\omega)).
$$

For each $v \in \mathcal{W}$, the same construction applies to the probabilistic future $(E^v, \mathbb{P}^v)$. This defines a collection of measurable maps

$$
H^v : \Omega^v \to \mathbb{R}, \tag{51}
$$

with $H^\emptyset = G$. By construction, the family $H = (H^v)_{v \in \mathcal{W}}$ is a homogeneous function (Definition 7.2). According to point 2 of Lemma 7.7 below, H satisfies:

$$
\forall v \in \mathcal{W}, \quad v \subset \omega \Rightarrow H^v(\omega \ominus v) = H^\emptyset(\omega), \quad \text{with probability 1.}
$$

Combined with Theorem 7.1, this implies that $H^\emptyset$ is constant on Ω. Let $\alpha(\mathbf{s}_0)$ denote this constant. Then, according to point 1 of Lemma 7.7 below, for every sequence $(W_n)_{n \geq 0}$ of stopping operators, we have:

$$
lim_{n \to \infty} M(f_0, W_n) = H^\emptyset = \alpha(\mathbf{s}_0), \quad \text{with probability 1.}
$$

This shows that the ergodic mens $M(f, \cdot)$ converge to the constant $\alpha(\mathbf{s}_0)$. $\qquad\square$

With Lemma 7.5, we have shown points 1 and 2 of Theorem 6.1. It remains to show point 3, which is the topic of next lemma.

Lemma 7.6 *The density coefficients $\alpha(\mathbf{s})$ satisfy:*

$$
\forall \mathbf{s} \in \Sigma, \quad 0 < \alpha(\mathbf{s}) \leq 1, \qquad and \qquad \sum_{\mathbf{s} \in \Sigma} \alpha(\mathbf{s}) = 1. \tag{52}
$$

Proof. We first show that the $\alpha(\mathbf{s})$'s sum to 1. Consider any regular sequence of stopping operators $(V_n)_{n\geq 0}$—we have seen the existence of such a sequence in Lemma 6.1. We have obviously:

$$\forall n \geq 0, \quad \sum_{\mathbf{s}\in\Sigma}\langle \mathbf{1_s}, V_n\rangle = \langle \mathbf{1}, V_n\rangle,$$

where $\mathbf{1}$ is the extended state function that counts all branching cells. Therefore, taking the ratio and then the limit, we get:

$$1 = \lim_{n\to\infty}\frac{1}{\langle\mathbf{1}, V_n\rangle}\sum_{\mathbf{s}\in\Sigma}\langle\mathbf{1_s}, V_n\rangle = \lim_{n\to\infty}\sum_{\mathbf{s}\in\Sigma}M(\mathbf{1_s}, V_n) = \sum_{\mathbf{s}\in\Sigma}\alpha(\mathbf{s}).$$

This shows that the $\alpha(\mathbf{s})$'s sum to 1, as claimed. This also implies that $0 \leq \alpha(\mathbf{s}) \leq 1$ for every $\mathbf{s}\in\Sigma$.

It remains only to show that $\alpha(\mathbf{s}) > 0$ for every $\mathbf{s}\in\Sigma$. Let $\mathbf{s}\in\Sigma$, and consider the sequence $(S_n^{\mathbf{s}})_{n\geq 1}$ of local returns to $\mathbf{s}$. This sequence is a regular sequence of stopping operators according to Proposition 7.1, §7.3. We have thus, applying the LLN to the extended state function $\mathbf{1_s}$:

$$\alpha(\mathbf{s}) = \lim_{n\to\infty}\frac{1}{\langle\mathbf{1}, S_n^{\mathbf{s}}\rangle}\langle\mathbf{1_s}, S_n^{\mathbf{s}}\rangle = \lim_{n\to\infty}\frac{n}{\langle\mathbf{1}, S_n^{\mathbf{s}}\rangle}. \tag{53}$$

The latter ratio can be written as follows:

$$\frac{n}{\langle\mathbf{1}, S_n^{\mathbf{s}}(\omega)\rangle} = \frac{n}{\langle\mathbf{1}, S_1^{\mathbf{s}}(\omega)\rangle + \sum_{k=1}^{n-1}\left(\langle\mathbf{1}, S_{k+1}^{\mathbf{s}}(\omega)\rangle - \langle\mathbf{1}, S_k^{\mathbf{s}}(\omega)\rangle\right)}.$$

We have for every $k \geq 1$:

$$\langle\mathbf{1}, S_{k+1}^{\mathbf{s}}\rangle - \langle\mathbf{1}, S_k^{\mathbf{s}}\rangle = \langle\mathbf{1}, S_{k+1}^{\mathbf{s}} \ominus S_k^{\mathbf{s}}\rangle = \langle\mathbf{1}, S^{\mathbf{s}, m_k}\rangle, \tag{54}$$

where m_k denotes the marking $\gamma(S_k^{\mathbf{s}})$, and $S^{\mathbf{s}, m}$ denotes as in (37) the local return to $\mathbf{s}$ defined on Ω_{E^m}. It follows from the Markov property for Markov nets [2] that the sequence $(S^{\mathbf{s}, m_k})_{k\geq 1}$ is a sequence of independent random variables. They are not identically distributed since the law of $S^{\mathbf{s}, m_k}$ depends on $m_k = \gamma(S_k^{\mathbf{s}})$. But these laws range over a finite set, since markings are finitely many. Furthermore, each expectation $\mathbb{E}(\langle\mathbf{1}, S^{\mathbf{s}, m}\rangle)$ is finite according to Lemma 7.4. Therefore, we shall apply Proposition A.1 of Appendix A to get, using (53) and (54):

$$\frac{1}{\alpha(\mathbf{s})} \leq \max_m \mathbb{E}^m\langle\mathbf{1}, S^{\mathbf{s}, m}\rangle < \infty.$$

This shows that $\alpha(\mathbf{s}) > 0$. $\qquad\square$

With Lemmas 7.5 and 7.6, we have completed the proof of Theorem 6.1. In the proof of Lemma 7.5, we have used the following lemma:

Lemma 7.7 *Assume that $\mathcal{N}$ is a recurrent Markov net with integrable concurrency height, and let H be the homogeneous function defined in (51). Consider the state function $f_0 = \mathbf{1}_{\mathbf{s}_0}$ defined by (50). Then H satisfies:*

1. For every regular sequence $(W_n)_{n \geq 0}$ of stopping operators, we have:

$$\lim_{n \to \infty} M(f_0, W_n) = H^{\emptyset}, \quad \text{with probability } 1.$$

2. For each $v \in \mathcal{W}$, and with probability 1:

$$v \subseteq \omega \Rightarrow H^v(\omega \ominus v) = H^{\emptyset}(\omega).$$

Proof. Let $(W_n)_{n \geq 0}$ be a regular sequence of stopping operators. $(V_j, Z_{j+1})_{j \geq 0}$ denotes as above the normal decomposition of maximal configurations. For each $n \geq 1$, let $J_n : \Omega \to \mathbb{N}$ be the integer-valued random variable defined by:

$$J_n(\omega) = \inf\{p \geq 0 : V_p(\omega) \supseteq W_n(\omega)\}. < \infty, \quad \text{with probability } 1.$$

Then, with probability 1, $\delta_E(V_{J_n}(\omega)) \cap \delta_E(W_n(\omega)) \neq \emptyset$. Therefore $E^{W_n(\omega)}$ and $E^{V_{J_n}(\omega)}$ possess a minimal event in common. Thus $V_{J_n}(\omega)$ is a sub-configuration of ω that keeps a token frozen, starting from the sub-configuration $W_n(\omega)$. Therefore, by definition of the integer random variables $L_m(b, \cdot)$, we have:

$$\langle N, V_{J_n} \ominus W_n \rangle \leq \Gamma_{\gamma(W_n)}, \tag{55}$$

where $\Gamma_m(\xi) = \sup_{b \in m} L_m(b, \xi)$ for all recursively stopped markings m and $\xi \in \Omega_{E^m}$. Remark that, since all $L_m(b, \cdot)$ are integrable, and since there are only finitely many of them, $\Gamma = (\Gamma_m)_{m \in \mathbf{M}_{\mathrm{rs}}}$ is a homogeneous integrable function.

We now show that the quantity:

$$\epsilon_n = M(f_0, W_n) - \frac{\langle f_0, V_{J_n} \rangle}{\langle N, V_{J_n} \rangle}$$

goes to 0 when n goes to ∞, with probability 1.

$$\begin{aligned}
\epsilon_n &= \frac{\langle f_0, W_n \rangle}{\langle N, W_n \rangle} - \frac{\langle f_0, V_{J_n} \rangle}{\langle N, V_{J_n} \rangle} \\
&= \frac{\langle f_0, W_n \rangle - \langle f_0, V_{J_n} \rangle}{\langle N, W_n \rangle} + \langle f_0, V_{J_n} \rangle \left(\frac{1}{\langle N, W_n \rangle} - \frac{1}{\langle N, V_{J_n} \rangle} \right) \\
&= \frac{-\langle f_0, V_{J_n} \ominus W_n \rangle}{\langle N, W_n \rangle} + \frac{\langle f_0, V_{J_n} \rangle}{\langle N, V_{J_n} \rangle} \frac{\langle N, V_{J_n} \ominus W_n \rangle}{\langle N, W_n \rangle}.
\end{aligned}$$

We use that $\langle f_0, W_n \rangle \leq \langle N, W_n \rangle$ for every $n \geq 1$ to get:

$$|\epsilon_n| \leq 2 \frac{\langle N, V_{J_n} \ominus W_n \rangle}{\langle N, W_n \rangle} \leq 2 \frac{\Gamma_{\gamma(W_n)}(\omega \ominus W_n)}{\langle N, W_n \rangle}, \tag{56}$$

where the latter inequality follows from (55). According to Proposition 7.2, and since $(\Gamma_m)_{m \in \mathbf{M}_{\mathrm{rs}}}$ is an integrable homogeneous function, the right member in (56) goes to 0 with probability 1, and thus $\lim_{n \to \infty} \epsilon_n = 0$ with probability 1. This completes the proof of point 1 of Lemma 7.7.

The proof of point 2 follows from a similar calculation. $\square$

7.5.2 Proof of Theorem 6.2

We shall now prove Theorem 6.2 as a corollary of Theorem 6.1. We also use the result stated in Lemma 7.3, that uses the fact that we consider distributed probabilities.

Let $f = (f(\mathbf{s}, \cdot))_{\mathbf{s} \in \Sigma}$ be any extended state function; we have to study the convergence and the limit of the ergodic means $M(f, V_n)$, for $(V_n)_{n \geq 1}$ any regular sequence of stopping operators. Because of the linearity of $f \to M(f, \cdot)$, we assume without loss of generality that f satisfies $f(\mathbf{s}, \cdot) = 0$ on $\Omega_{\mathbf{s}}$ whenever $\mathbf{s} \neq \mathbf{s}_0$, where $\mathbf{s}_0$ is some particular dynamic cluster. Let $(V_n)_{n \geq 0}$ be a regular sequence of stopping operators, and let $\omega \in \Omega$ such that $\lim_{n \to \infty} M(\mathbf{1}_{\mathbf{s}_0}, V_n(\omega)) = \alpha(\mathbf{s}_0)$ holds. Let also $w_n = S_n^{\mathbf{s}_0}(\omega)$ denote the sequence of local return operators to $\mathbf{s}_0$ applied to ω. According to Proposition 7.1, we assume without loss of generality that $\sup_{n \geq 1} w_n = \omega$, since such ω have probability 1.

Recall that the branching cells $x \in \Delta(\omega)$ such that $\langle x \rangle = \mathbf{s}_0$ are totally ordered. Therefore, for every $n \geq 1$, the set of branching cells:

$$I_n = \{x \in \Delta(V_n(\omega)) \ : \ \langle x \rangle = \mathbf{s}_0\}$$

is an interval. But since $\sup_{n \geq 1} w_n = \omega$, there is thus an integer $J(n)$ such that:

$$I_n = \{x \in \Delta(w_{J(n)}) \ : \ \langle x \rangle = \mathbf{s}_0\}.$$

Since we assume that $f(\mathbf{s}, \cdot)$ vanishes if $\mathbf{s} \neq \mathbf{s}_0$, we have:

$$
\begin{aligned}
M(f, V_n) &= \frac{1}{\langle \mathbf{1}, V_n \rangle} \langle f, V_n \rangle \\
&= \frac{1}{\langle \mathbf{1}, V_n \rangle} \langle f, w_{J(n)} \rangle \\
&= \frac{\langle \mathbf{1}_{\mathbf{s}_0}, V_n \rangle}{\langle \mathbf{1}, V_n \rangle} \frac{1}{\langle \mathbf{1}_{\mathbf{s}_0}, V_n \rangle} \sum_{k=1}^{J(n)} f(\mathbf{s}_0, Y_k^{\mathbf{s}_0}),
\end{aligned}
\tag{57}
$$

where $Y_k^{\mathbf{s}_0}$ is defined by (39). $\Delta(\omega)$ contains infinitely many instances of $\mathbf{s}_0$ since the net is recurrent. Since $\sup_n V_n(\omega) = \omega$, it follows that $\lim_{n \to \infty} J(n) = +\infty$. The sequence $(Y_n^{\mathbf{s}_0})_{n \geq 1}$ is a sequence of independent identically distributed (i.i.d.) random variables of law $p_{\mathbf{s}_0}$ according to Lemma 7.3. It follows thus from the Strong law of large numbers for

i.i.d. sequences (see Theorem A.2 in Appendix A) that we have:

$$\frac{1}{\langle \mathbf{1}_{\mathbf{s}_0}, V_n \rangle} \sum_{k=1}^{J(n)} f(\mathbf{s}_0, Y_k^{\mathbf{s}_0}) = \frac{1}{\langle \mathbf{1}_{\mathbf{s}_0}, w_{J(n)} \rangle} \sum_{k=1}^{J(n)} f(\mathbf{s}_0, Y_k^{\mathbf{s}_0})$$

$$= \frac{1}{J(n)} \sum_{k=1}^{J(n)} f(\mathbf{s}_0, Y_k^{\mathbf{s}_0}) \longrightarrow_{n \to \infty} p_{\mathbf{s}_0}(f), \tag{58}$$

where $p_{\mathbf{s}_0}(f)$ is defined by:

$$p_{\mathbf{s}_0}(f) = \sum_{z \in \Omega_{\mathbf{s}_0}} p_{\mathbf{s}_0}(z) f(\mathbf{s}_0, z).$$

On the other hand, we have

$$\lim_{n \to \infty} \frac{\langle \mathbf{1}_{\mathbf{s}_0}, V_n \rangle}{\langle \mathbf{1}, V_n \rangle} = \alpha(\mathbf{s}_0). \tag{59}$$

Combining (57), (58) and (59), we get:

$$\lim_{n \to \infty} M(f, V_n) = \alpha(\mathbf{s}_0) p_{\mathbf{s}_0}(f),$$

which is Theorem 6.2 for our particular f. As it was enough to consider this particular f, the proof of the theorem is complete.

8 Summary and Perspectives

We have proposed branching cells and dynamic clusters as a notion of local, concurrent, state for event structures and Petri nets. We have applied the construction of distributed probabilities for event structures to the randomization of traces of safe Petri nets. This results in the model of Markov nets, a proper generalization of Markov chains to true-concurrency systems. The Law of Large Numbers extends to Markov nets, with dynamic clusters taken as states.

Although our work relies on the local finiteness assumption for event structures, the notion of branching cells that we have developed can be applied to general event structures arising from Petri net unfoldings. Some new issues arise for non-locally finite event structures, however. First, branching cells may be infinite; second, maximal configurations may not be R-stopped, hence the randomization is not immediate as for locally finite event structures. We are currently working on this extended setting for Markov nets.

A Appendix: Classical Laws of Large Numbers

In this section we state the classical Laws of Large Numbers that we use, for Markov chains and for sequences of independent and sequences of independent identically distributed (i.i.d.) random variables.

Theorem A.1 [9, Ch.V, Th. 6.1 p.219] *Let $(X_n)_{n \geq 0}$ be a Markov chain on a finite state space Q, and let $f : Q \to \mathbb{R}$ be a real-valued function. Then the ergodic means:*

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k)$$

have a limit, for $n \to \infty$, with probability 1.

Theorem A.2 [9, Ch.III, Th. 5.1 p.142] *Let $(X_n)_{n \geq 0}$ be a sequence of i.i.d. random variables, with probability law p on a countable set Q. For any real-valued nonnegative function $f : Q \to \mathbb{R}$, the following limit holds with probability 1:*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \sum_{s \in Q} p(s) f(s).$$

Finally, the following result is an exercise:

Proposition A.1 *Let $P = \{p_i, \, i = 1, \ldots, r\}$ be a finite family of probability laws on some countable set Q. Let $(X_n)_{n \geq 0}$ be a sequence of independent random variables, such that, for each $n \geq 0$, the law of X_n belongs to P. Let $f : Q \to \mathbb{R}$ be a nonnegative real-valued function, such that, for each $p_i \in P$, the following expectation is finite:*

$$q_i(f) = \sum_{q \in Q} p_i(q) f(q) < \infty.$$

Then we have:

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \leq \max_{i=1,\ldots,r} q_i(f) < \infty.$$

References

[1] S. Abbes. *Probabilistic Model for Concurrent and Distributed Systems. Limit Theorems and Application to Statistical Parametric Estimation.* PhD thesis, IRISA/Université de Rennes 1, 2004. Available from `ftp://ftp.irisa.fr/techreports/theses/2004/abbes.pdf`.

[2] S. Abbes. A (true) concurrent Markov property and some applications to Markov nets. In G. Ciardo and P. Darondeau, editors, *Proc. of ICATPN 2005*, volume 3536 of *LNCS*, pages 70–89, 2005.

[3] S. Abbes and A. Benveniste. Probabilistic true-concurrency models: Branching cells and distributed probabilities for event structures. To appear in *Information & Computation.*

[4] S. Abbes and A. Benveniste. Branching cells as local states for event structures and nets: probabilistic applications. Technical Report RR-5347, INRIA, 2004. Available at `http://www.inria.fr/rrrt/rr-5347.html`.

[5] S. Abbes and A. Benveniste. Branching cells as local states for event structures and nets: probabilistic applications. In V. Sassone, editor, *Proc. of FOSSACS 2005*, volume 3441 of *LNCS*, pages 95–109, 2005.

[6] A. Benveniste, S. Haar, and E. Fabre. Markov nets: probabilistic models for distributed and concurrent systems. *IEEE Transactions on Automatic Control*, 48(11):1936–1950, 2003.

[7] P. Billingsley. *Probability and Measure.* John Wiley & Sons, 1986. Second edition.

[8] J. Desel and J. Esparza. *Free choice Petri nets.* Cambridge University Press, 1995.

[9] J.L. Doob. *Stochastic Processes.* John Wiley & Sons, 1953.

[10] A. Mazurkiewicz. Trace theory. In W. Brauer, W. Reisig, and G. Rozenberg, editors, *Petri Nets: Applications and Relationships to Other Models of Concurrency*, volume 255 of *LNCS*, pages 279–324. Springer, 1986.

[11] P. Nielsen, G. Plotkin, and G. Winskel. Petri nets, event structures and domains, part 1. *Theoretical Computer Science*, 13:85–108, 1981.

[12] A. Raugi and J.-P. Conze. Cours de DEA. Martingales et systèmes dynamiques. IRMAR, Université de Rennes 1, 2002.

[13] D. Revuz. *Markov Chains.* North Holland, 1975.

[14] D. Varacca, H. Völzer, and G. Winskel. Probabilistic event structures and domains. In P. Gardner and N. Yoshida, editors, *Proc. of CONCUR 2004*, volume 3170 of *LNCS*, pages 481–496, 2004.

[15] H. Völzer. Randomized non-sequential processes. In K.G. Larsen and M. Nielsen, editors, *Proc. of CONCUR 2001*, volume 2154 of *LNCS*, pages 184–201, 2001.