Basic Elements for the Performance Evaluation of Networks (PEN)

Master 2 UBINET – SI5
Université Côte d'Azur
Academic Year 2024-2025

Sara Alouf
Sara.Alouf@inria.fr

based on the lecture notes of Philippe Nain

Contents

Foreword

The objective of the course is to introduce analytical tools that are needed to construct/measure/analyze models of resource contention systems, e.g. computer systems, networks. This course enables one to answer questions such as:

- what is the throughput of the wifi 802.11 protocol?
- if both the arrival and the service rates double, will the response time stay the same?
- given the choice between a single machine with speed s, or n machines each with speed s/n, which should one choose?
- how many operators in a call center are needed to keep the rejection probability of incoming requests low?

Given a system (Internet, department WLAN, web server, ...), we could be interested to know how well it performs, we could need to adjust its design or to change its scale. Appropriate performance metrics are the average response time, the quality of service (user's point of view), the throughput or the number of supported sessions (system's point of view). Other metrics are related to the utilization of the system, its reliability, or its availability.

The performance evaluation of networks/systems may be done using either *measurements* (by measuring the performance of existing systems), or *simulations* (this requires to build a software emulator of the system; then execute it; then use traces or random numbers to generate workload), or *analysis* (this requires to build a mathematical model that captures the essence of the system, then use mathematical tools to evaluate the performance), or a combinations of these methods.

	measurement	simulation	analysis
when	prototyping, configuration tuning, monitoring	anytime	anytime
effort	varies	moderate	low
cost	high	moderate	low
tools	instrumentation	languages	mathematics
accuracy	varies	moderate	low
salability	high	medium	low
scalability	low	medium	high

About the course

The course prerequisites are a good knowledge in probability theory and in linear algebra.

The course will cover Markov chains in 3 lectures and queueing theory in 3 lectures. The last session is devoted to few use case applications.

Recommended reading material: Mor Harchol Balter, “Performance Modeling and Design of Computer Systems”.

Course’s website: `http://www-sop.inria.fr/members/Sara.Alouf/PEN/`.
On Moodle: `https://lms.univ-cotedazur.fr/2024/course/view.php?id=14869`.

Course requirements: one homework for each of the first six lectures contributing 10% to the final grade; a final assessment in the form of a three hours practise session covering all seven lectures contributing 40% to the final grade.

Nota bene: homeworks are a *personal* effort. Copied solutions will get 0 for a grade.

About the lecture notes

The material presented in Chapters 1, 2 and 4 comes mainly from reference [5] below. Chapter 3 has been borrowed from [6]. Section 5.1 comes from [1]. Section 6.1 comes from [2]. Section 6.2 is taken from [4]. A highly approachable book covering most material presented here is [3].

References

[1] D. P. Bertsekas and R. G. Gallager, *Data Networks*, (2nd edition) Prentice Hall, 1992.

[2] E. Gelenbe and I. Mitrani, *Analysis and Synthesis of Computer Systems*, Academic Press (London and New York), 1980.

[3] M. Harchol-Balter, *Performance Modeling and Design of Computer Systems*, Cambridge University Press, 2013.

[4] F. P. Kelly, *Reversibility and Stochastic Networks*, Wiley, Chichester, 1979.

[5] L. Kleinrock, *Queueing Theory*, Vol. 1, J. Wiley & Sons, New York, 1975.

[6] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*, John Hopkins University Press, 1981.

Brief probability refresher

A probability refresher is available in Appendix A. The *sample-space* Ω is the set of all *outcomes* associated with an experiment.
To compute the probability that the event A occurs given that the event B has occurred, use the *Bayes formula*:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

The *law of total probability* states that if $A_1, A_2, \ldots, A_n$ are events such that

(a) $A_i \cap A_j = \emptyset$ if $i \neq j$ (mutually exclusive events)

(b) $P(A_i > 0)$ for $i = 1, 2, \ldots, n$

(c) $A_1 \cup A_2 \cup \cdots \cup A_n = \Omega$.

then, for any event A, $P(A) = \sum_{i=1}^{n} P(A \cap A_i) = \sum_{i=1}^{n} P(A \mid A_i)\, P(A_i)$.

A collection of random variables (rvs) $\mathbf{X} = \{X(t),\, t \in T\}$ is called a *stochastic process* (see Appendix B). In other word, for each $t \in T$, $X(t)$ is a rv mapping from Ω into some set $E \subset \mathbb{R}$ (e.g., $E = [0, \infty)$, $E = \mathbb{N}$) with the interpretation that $X(t)(\omega)$ (also written as $X(t, \omega)$) is the value of the stochastic process $\mathbf{X}$ at time t on the outcome (or path) ω.

The space E is called the *state-space* of the stochastic process $\mathbf{X}$. If the set E is countable then $\mathbf{X}$ is called a *discrete-space* stochastic process; if the set E is continuous then $\mathbf{X}$ is called a *continuous-space* stochastic process.

If T is countable (e.g., $T = \mathbb{N}$, $T = \{\ldots, -2, -1, 0, 1, 2 \ldots\}$) then $\mathbf{X}$ is called a *discrete-time* stochastic process; if T is continuous (e.g., $T = \mathbb{R}$, $T = [0, \infty)$) then $\mathbf{X}$ is called a *continuous-time* stochastic process. When T is countable one will in general substitute the notation $X(t)$ for $X(n)$.

The *Poisson process* (see Appendix C) is an event counting process. A Poisson process $(N(t),\, t \in T)$ having rate λ is such that (*i*) $N(0) = 0$, (*ii*) the increments of $N(t)$ are independent, and (*iii*) the number of events in any interval t is a Poisson distributed random variable, i.e.

$$P\big(N(t+s) - N(s) = k\big) = \frac{e^{-\lambda t} (\lambda t)^k}{k!} \qquad k = 0, 1, \ldots$$

The inter-event durations of a Poisson process with rate λ are iid and follow an exponential distribution with parameter λ.

Part I: Markov Chains

A *Markov*[1] *process* is a stochastic process $(X(t),\, t \in T)$, $X(t) \in E \subset \mathbb{R}$, such that

$$P(X(t) \le x \mid X(t_1) = x_1, \ldots, X(t_n) = x_n) = P(X(t) \le x \mid X(t_n) = x_n) \tag{1}$$

for all $x_1, \ldots, x_n, x \in \mathcal{E}$, $t_1, \ldots, t_n, t \in T$ with $t_1 < t_2 < \cdots < t_n < t$.

Intuitively, (1) says that the *probabilistic future* of the process depends only on the *current* state and not upon the *history* of the process. In other words, the entire history of the process is summarized in the current state.

In the following we shall be mostly concerned with discrete-space Markov processes, commonly referred to as *Markov chains*.

We shall distinguish between discrete-time Markov chains and continuous-time Markov chains.

1 Discrete-Time Markov Chain

A discrete-time Markov Chain (abbreviated as DTMC) is a discrete-time (with index set $\mathbb{N} := \{0, 1, \ldots\}$) discrete-space (with state-space $\mathcal{E}$) stochastic process $(X(n),\, n \in \mathbb{N})$ such that for all $n \ge 0$

$$P(X(n+1) = j \mid X(0) = i_0, X(1) = i_1, \ldots, X(n-1) = i_{n-1}, X(n) = i) = P(X(n+1) = j \mid X(n) = i) \tag{2}$$

for all $i_0, \ldots, i_{n-1}, i, j \in \mathcal{E}$. Equation (2) is called the *Markov property*.

A DTMC is called a *finite-state* DTMC if the set $\mathcal{E}$ is finite.

A DTMC is *homogeneous* if $P(X(n+1) = j \mid X(n) = i)$ does not depend on n for all $i, j \in \mathcal{E}$. If so, we shall write

$$p_{i,j} = P(X(n+1) = j \mid X(n) = i) \quad \forall i, j \in \mathcal{E}.$$

$p_{i,j}$ is the *one-step transition probability* from state i to state j, that is the probability to go from state i to state j in exactly one time-step.

Unless otherwise mentioned we shall only consider homogeneous DTMC's.

Define $\mathbf{P} = [p_{i,j}]$ to be the transition matrix of a DTMC, namely,

$$\mathbf{P} = \begin{pmatrix} p_{0,0} & p_{0,1} & \cdots & p_{0,j} & \cdots \\ p_{1,0} & p_{1,1} & \cdots & p_{1,j} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{i,0} & p_{i,1} & \cdots & p_{i,j} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \tag{3}$$

[1] Andrey Andreyevitch Markov was a Russian mathematician.

We must have

$$\begin{aligned} p_{i,j} &\geq 0 \quad \forall i,j \in \mathcal{E} &(4)\\ \sum_{j\in\mathcal{E}} p_{i,j} &= 1 \quad \forall i \in \mathcal{E}. &(5)\end{aligned}$$

Equation (5) is a consequence of axiom (b) of a probability measure (see Section A.1 and says that the sum of the elements in each row is 1. A matrix satisfying (4) and (5) is called a *stochastic matrix*.

If the state-space $\mathcal{E}$ is finite (say, with k states) then $\mathbf{P}$ is a k-by-k matrix; otherwise $\mathbf{P}$ has infinite dimension.

Example 1. Consider a sequence of Bernoulli trials in which the probability of success (S) on each trial is p and that of failure (F) is q, where $p+q=1$, $0<p<1$. Let the state of the process at trial n (i.e., $X(n)$, with $n \geq 0$) be the number of uninterrupted successes that have been completed at this point, so that $\mathcal{E} = \mathbb{N}$. For instance, if the first 5 outcomes where SFSSF then $X(0)=1$, $X(1)=0$, $X(2)=1$, $X(3)=2$ and $X(4)=0$. The transition matrix is given by

$$\mathbf{P} = \begin{pmatrix} q & p & 0 & 0 & 0 & \dots \\ q & 0 & p & 0 & 0 & \dots \\ q & 0 & 0 & p & 0 & \dots \\ q & 0 & 0 & 0 & p & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

The state 0 can be reached in one transition from any state while the state $i+1$, $i \geq 0$, can only be reached from the state i (with probability p). Clearly, this DTMC is homogeneous. ♦

We now define the n-step transition probabilities $p_{i,j}^{(n)}$ by

$$p_{i,j}^{(n)} = P(X(n) = j \,|\, X(0) = i) \tag{6}$$

for all $i,j \in \mathcal{E}$, $n \geq 0$. $p_{ij}^{(n)}$ is the probability of going from state i to state j in n steps.

Proposition 1 (Chapman-Kolmogorov equation)**.** *For all $n \geq 0$, $m \geq 0$, $i,j \in \mathcal{E}$, we have*

$$p_{i,j}^{(n+m)} = \sum_{k\in\mathcal{E}} p_{i,k}^{(n)}\, p_{k,j}^{(m)} \tag{7}$$

or, in matrix notation,

$$\mathbf{P}^{(n+m)} = \mathbf{P}^{(n)}\,\mathbf{P}^{(m)} \tag{8}$$

where $\mathbf{P}^{(n)} := \left[p_{i,j}^{(n)}\right]$. Therefore,

$$\mathbf{P}^{(n)} = \mathbf{P}^{n} \quad \forall n \geq 1 \tag{9}$$

where $\mathbf{P}^n$ is the n-th power of the matrix $\mathbf{P}$. ■

The Chapman-Kolmogorov equation merely says that if we are to travel from state i to state j in $n+m$ steps then we must do so by first traveling from state i to *some* state k in n steps and then from state k to state j in m more steps.

Proof. The proof of Proposition 1 goes as follows. We have

$$\begin{aligned}
p_{i,j}^{(n+m)} &= P(X(n+m)=j \,|\, X(0)=i) \qquad \text{(use now law of total probability)} \\
&= \sum_{k\in\mathcal{E}} P(X(n+m)=j \,|\, X(0)=i, X(n)=k) \times P(X(n)=k \,|\, X(0)=i) \\
&= \sum_{k\in\mathcal{E}} P(X(n+m)=j \,|\, X(n)=k)\, P(X(n)=k \,|\, X(0)=i) \qquad \text{(Markov property (2))} \\
&= \sum_{k\in\mathcal{E}} p_{k,j}^{(m)}\, p_{i,k}^{(n)}
\end{aligned}$$

since the DTMC is homogeneous, which proves (7) and (8).

Let us now establish (9). Since $\mathbf{P}^{(1)} = \mathbf{P}$, we see from (8) that $\mathbf{P}^{(2)} = \mathbf{P}^2$, $\mathbf{P}^{(3)} = \mathbf{P}^{(2)}\,\mathbf{P} = \mathbf{P}^3$ and, more generally, that $\mathbf{P}^{(n)} = \mathbf{P}^n$ for all $n \ge 1$. This concludes the proof. ★

Example 2. Consider a communication system that transmits the digits 0 and 1 through several stages. At each stage, the probability that the same digit will be received by the next stage is 0.75. What is the probability that a 0 entered at the first stage is received as a 0 after the fifth stage?

We want to find $p_{0,0}^{(5)}$ for a DTMC with transition matrix $\mathbf{P}$ given by

$$\mathbf{P} = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

From Proposition 1 we know that $p_{0,0}^{(5)}$ is the $(1,1)$-entry of the matrix $\mathbf{P}^5$. We can compute $\mathbf{P}^2$, then $\mathbf{P}^4$, then $\mathbf{P}^5$:

$$\mathbf{P}^2 = \frac{1}{8}\begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}, \quad \mathbf{P}^4 = \frac{1}{32}\begin{bmatrix} 17 & 15 \\ 15 & 17 \end{bmatrix}, \quad \mathbf{P}^5 = \frac{1}{64}\begin{bmatrix} 33 & 31 \\ 31 & 33 \end{bmatrix}.$$

We find $p_{0,0}^{(5)} = 33/64 = 0.515625$. ♦

So far, we have only been dealing with conditional probabilities. For instance, $p_{i,j}^{(n)}$ is the probability that the system in state j at time n given it was in state i at time 0. We have shown in Proposition 1 that this probability is given by the (i,j)-entry of the power matrix $\mathbf{P}^n$. What we would like to do now is to compute the unconditional probability that the system is in state i at time n, namely, we would like to compute $\pi_i(n) := P(X(n)=i)$.

This quantity can only be computed if we provide the initial distribution function of $X(0)$, that is, if we provide $\pi_i(0) = P(X(0)=i)$ for all $i \in \mathcal{E}$, where of course $\sum_{i\in\mathcal{E}} \pi_i(0) = 1$.

In that case, from the law of total probability, we can write

$$\begin{aligned}
P(X(n)=j) &= \sum_{i\in\mathcal{E}} P(X(n)=j \,|\, X(0)=i)\, \pi_i(0) \\
&= \sum_{i\in\mathcal{E}} p_{i,j}^{(n)}\, \pi_i(0)
\end{aligned}$$

from Proposition 1 or, equivalently, in matrix notation:

Proposition 2. *For all* $n \geq 1$,

$$\pi(n) = \pi(0)\,\mathbf{P}^n. \tag{10}$$

where $\pi(m) := (\pi_0(m), \pi_1(m), \ldots)$ *for all* $m \geq 0$*. From* (10) *we deduce that (one can also obtain this result directly)*

$$\pi(n+1) = \pi(n)\,\mathbf{P} \quad \forall n \geq 0. \tag{11}$$

■

Assume that the limiting state distribution function $\lim_{n\to\infty} \pi_i(n)$ exists for all $i \in \mathcal{E}$. Call it π_i and let $\pi = (\pi_i, i \in \mathcal{E})$. Observe that taking the limit as $n \to \infty$ in (11) yields

$$\pi = \pi\,\mathbf{P} \qquad \Leftrightarrow \qquad \pi_i = \sum_{j\in\mathcal{E}} \pi_j p_{j,i}\ , \quad i \in \mathcal{E}.$$

How can one compute π? A natural answer is "by solving the system of linear equations defined by $\pi = \pi\,\mathbf{P}$" to which one should add the normalization condition $\sum_{i\in\mathcal{E}} \pi_i = 1$ or, in matrix notation, $\pi\,\mathbf{1} = 1$, where $\mathbf{1}$ is the column vector where every component is 1.

We shall now give conditions under which the above results hold (i.e., $(\pi_0(n), \pi_1(n), \ldots)$ has a limit as n goes to infinity and this limit solves the system of equations $\pi = \pi\,\mathbf{P}$ and $\pi\,\mathbf{1} = 1$).

To do so, we need to introduce several notions.

For every state $i \in \mathcal{E}$, define the integer $d(i)$ as the greatest common divisor of all integers n such that $p_{i,i}^{(n)} > 0$. $d(i) = \gcd\{n | p_{i,i}^{(n)} > 0\}$. If $d(i) = 1$ then the state i is *aperiodic*.

A DTMC chain is *aperiodic* if all states are aperiodic.

There is also the notion of *communication* between the states. We shall say that a state j is *reachable* from a state i if $p_{i,j}^{(n)} > 0$ for some $n \geq 1$. If j is reachable from i and if i is reachable from j then we say that i and j communicate, and write $i \leftrightarrow j$.

A DTMC is *irreducible* if $i \leftrightarrow j$ for all $i, j \in \mathcal{E}$.

Another notion relates to the probability of returning to a given state. Let f_i be the probability that a chain starting in state i ever returns to state i. If $f_i = 1$ then state i is said to be *recurrent*; if $f_i < 1$ then state i is said to be *transient*. The number of visits to a recurrent state (respectively transient state) is infinite (respectively finite) with probability 1.

If the mean time between recurrences (returning to same state) is finite (respectively infinite) then the state is said to be *positive recurrent* (respectively *null recurrent*).

For a null recurrent state the number of visits to this state is infinite but at the same time the mean time between visits is also infinite!

If a state i is recurrent/transient and i and j communicate, then j is recurrent/transient. A direct consequence is that in an irreducible Markov chain, states are either all recurrent or all transient.

A DTMC is *positive recurrent* if all states are positive recurrent.

In a transient Markov chain the limiting distribution does not exist (the probability of being in a transient state is 0 after a large enough number of steps). A transient Markov chain has no stationary distribution.

Ergodicity: Let $S_j(n)$ be the total time spent in state j during the first n units of time. This is a rv that depends on the path taken by the Markov chain. For any j in the state-space $\mathcal{E}$, and for almost any path, we have

$$\lim_{n\to\infty} \frac{S_j(n)}{n} = \pi_j. \tag{12}$$

In other words, over a "long" period of time, the proportion of time spent in state j is close to the probability of finding the chain in state j at the end of the interval.

Ergodicity means that temporal averages and spatial averages coincide. An *ergodic* DTMC is one that has all three desirable properties: aperiodicity, irreducibility, and positive recurrence.

We have the following fundamental result of DTMC theory.

Proposition 3 (Invariant probability measure of a DTMC)**.** *If a homogeneous DTMC with transition matrix* $\mathbf{P}$ *is irreducible and aperiodic, and if the system of equations*

$$\begin{aligned} \mathbf{x} &= \mathbf{xP} \\ \mathbf{x}\cdot\mathbf{1} &= 1 \end{aligned}$$

has a unique strictly positive solution (i.e. for all $i \in \mathcal{E}$*,* $x(i)$*, the ith element of the row vector* $\mathbf{x}$*, is strictly positive) denoted by* $\pi = (\pi_j,\ j \in \mathcal{E})$*, then*

$$\pi_j = \lim_{n\to\infty} P(X(n) = j) \tag{13}$$

for all $i \in \mathcal{E}$*, regardless of the initial state* $X(0)$*.* ■

Observe that the convergence in (13) does not depend on the initial state i. In other words, for any initial state the probability distribution of $X(n)$ will converge to π as n goes to infinity.

We shall not prove this result. The equation $\pi = \pi\,\mathbf{P}$ is called the *invariant equation* and π is usually referred to as the *invariant probability*. Once a Markov chain has reached stationarity its probability distribution does not change.

Let us briefly comment on the conditions that the DTMC has to be irreducible and aperiodic for Proposition 3 to hold.

Irreducibility: consider a three-state DTMC (with states 1, 2 and 3) with transition matrix

$$\mathbf{P} = \begin{pmatrix} 0.3 & 0.7 & 0 \\ 0.1 & 0.9 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In words, if the chain is in state 3 at time 0 then it will stay in state 3 forever, while if at time 0 it is in state 1 or in state 2 then it will never leave these states. This DTMC is not irreducible since state 1 and 2 cannot be reached from state 3 and conversely. Therefore, it is obvious that the probability to find this DTMC in state 3 depends on whether or not the chain is in state 3 at time 0 (if then this probability is 1 and if not then it is 0).

A famous example of DTMC that is not irreducible is the one corresponding to the web graph. Tricks need to be made to make the chain irreducible so as to find the PageRank of each page (i.e. find the stationary distribution).

Aperiodicity: consider a two-state DTMC (with state 1 and 2) with transition matrix

$$\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In words, the chain alternates between state 1 and 2. Therefore, if the chain is in state 1 at time 0, namely $X(0) = 1$, then $P(X(n) = 2 \,|\, X(0) = 1) = 1$ if n is odd and $P(X(n) = 2 \,|\, X(0) = 1) = 0$ if n is an even integer. As a consequence $\lim_{n\to\infty} P(X(2n+1) = 2 \,|\, X(0) = 1) = 1$ and $\lim_{n\to\infty} P(X(2n) = 2 \,|\, X(0) = 1) = 0$ thereby showing that $P(X(n) = 2 \,|\, X(0) = 1)$ does not have a limit as n goes to infinity. The *limiting* distribution does not exist. This does not mean that the *stationary* distribution does not exist.

These two examples show the necessity for the chain to be irreducible and aperiodic if one wants Proposition 3 to hold.

Remark 1. *For an irreducible Markov chain having finite state-space, the system of equation*

$$\begin{aligned} \mathbf{x} &= \mathbf{xP} \\ \mathbf{x} \cdot \mathbf{1} &= 1 \end{aligned}$$

has always a unique strictly positive solution. This is the stationary *distribution, which coincides with the* limiting *distribution if the DTMC is aperiodic. For periodic DTMC, the* stationary *distribution represents the long-run time-average proportion of time spent in each state.*

When the state-space is infinite, it is not enough to have the irreducibility property for the stationary distribution to exist. What is needed is positive recurrence.

Example 3. *Consider an infinite DTMC with $\mathcal{E} = \{1, 2, \ldots\}$ having the following transition probabilities*

$$p_{1,1} = r; \quad p_{i,i} = s; \quad p_{i,i+1} = p; \quad p_{i+1,i} = q$$

for $i = 2, 3, \ldots$. We naturally have $r + p = 1$ and $p + q + s = 1$. The following can be verified

- *if $p < q$ then the chain is positive recurrent;*
- *if $p > q$ then the chain is transient;*
- *if $p = q$ (i.e. $s = 0$) then the chain is null recurrent.*

In the last two cases, the limiting probability at each state is 0 and both the limiting and the stationary distribution do not exist.

In practice one needs to check only the aperiodicity and the irreducibility of a DTMC. If the stationary equations have a solution (and this happens if the chain is positive recurrent), then this solution is the stationary and limiting distribution.

Remark 2. *The convergence of a DTMC to its stationary distribution is faster when the distance between the two largest eigen values of the transition matrix* $\mathbf{P}$ *is larger. Computing the stationary distribution using* (12) *becomes easier.*

1.1 Example of a discrete-time Markov chain

Consider a homogeneous discrete-time Markov chain (DTMC) on the state-space $\mathcal{E} = \{1, 2, 3\}$ with transition matrix $\mathbf{P}$ given by

$$\mathbf{P} = \begin{pmatrix} 0.2 & 0.2 & 0.6 \\ 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \end{pmatrix}. \tag{14}$$

One often represents a DTMC through its transition diagram. In this case it is given by

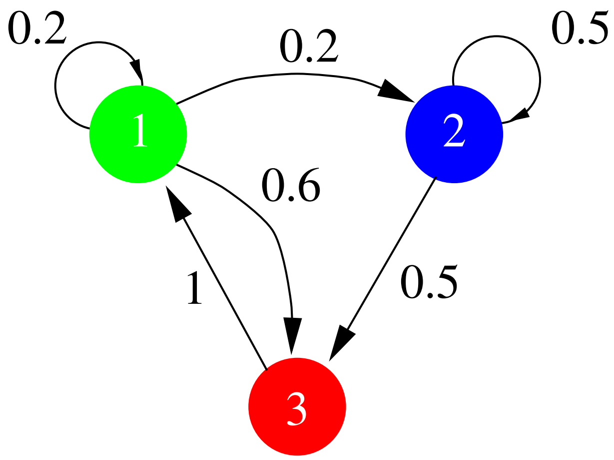

Figure 1: Transition diagram

The first type of question we may want to address is: *given that the chain is in state* 1 *at time* $t = 0$*, in what state will it be at time* $t = 4$*?* There is no deterministic answer to this question since, by construction, at time $t = 4$ the chain can a priori be in any of the states 1, 2 and 3. A more relevant question is: *given that the chain is in state* 1 *at time* $t = 0$*, what is the probability that it will be in state* j *at time* $t = 4$*?* The answer to the latter question is obtained using (10) in Proposition 2, which will allow us to compute $\pi(4) = (\pi_1(4), \pi_2(4), \pi_3(4))$, where we recall that is $\pi_i(n)$ the probability that the chain is in state $i \in \mathcal{E}$ at time $t = n$. Here, the initial probability distribution $\pi(0)$ is given by

$$\pi(0) = (1, 0, 0)$$

since we have specified that the chain is initially in state 1. We find

$$\pi(1) = (1, \quad 0, \quad 0) \begin{pmatrix} 0.2 & 0.2 & 0.6 \\ 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \end{pmatrix} = (0.2, 0.2, 0.6)$$

$$\pi(2) = (0.2, 0.2, 0.6) \begin{pmatrix} 0.2 & 0.2 & 0.6 \\ 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \end{pmatrix} = (0.64, 0.14, 0.22)$$

$$\pi(3) = (0.64, 0.14, 0.22) \begin{pmatrix} 0.2 & 0.2 & 0.6 \\ 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \end{pmatrix} = (0.348, 0.198, 0.454)$$

$$\pi(4) = (0.348, 0.198, 0.454) \begin{pmatrix} 0.2 & 0.2 & 0.6 \\ 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \end{pmatrix} = (0.5236, 0.1686, 0.3078).$$

In summary, at time $t = 4$, the chain will be

- in state 1 with probability 0.5236
- in state 2 with probability 0.1686
- in state 3 with probability 0.3078.

Let us now compute the stationary probabilities. First, note that the chain is clearly irreducible and aperiodic. Hence (cf. Proposition 3), if we find a unique strictly positive solution to the system of equations

$$\begin{cases} x_1 &= 0.2x_1 + 1.0x_3 \\ x_2 &= 0.2x_1 + 0.5x_2 \\ x_3 &= 0.6x_1 + 0.5x_2 \\ 1 &= x_1 + x_2 + x_3 \end{cases}$$

then this solution will be the stationary distribution of this Markov chain, that is, $\lim_{n\to\infty} \pi_i(n)$ will exist and will be given by x_i for $i \in \mathcal{E}$.

Solving for this system of equations gives

$$x = (5/11, 2/11, 4/11) \;\simeq\; (0.4545, 0.1818, 0.3636). \tag{15}$$

This solution being unique and strictly positive we conclude that this is the stationary probability of the chain, namely, $\lim_{n\to\infty} \pi_1(n) = 5/11$, $\lim_{n\to\infty} \pi_2(n) = 2/11$ and $\lim_{n\to\infty} \pi_3(n) = 4/11$.

In Figure 2 we have simulated the trajectory of the chain for 50 time units (starting in state 3 in red). The first bar is the state of the chain at $t = 1$, the second at $t = 2$, etc.

Note that the chain does not stay more that 1 time unit in state 3 (in red) which is normal since the probability of looping back to state 3 is zero (i.e. $p_{3,3} = 0$). More precisely, upon leaving state 3

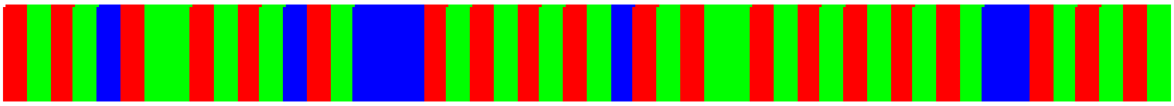

Figure 2: Trajectory after 50 units of time

(red) the chain always goes to state 1 (green) as $p_{3,1} = 1$ by definition of $\mathbf{P}$ in (14). We also observe that the state may stay several units of time in state 1 (green) and in state 2 (blue) but that it has the tendency to stay more often in state 2 (blue). This is also normal since $p_{2,2}$ is larger than $p_{3,3}$.

If we count the number of times that the chain has visited each state we find

22 times in state 1, 8 times in state 2, 20 times in state 3

that is, in percentage of time it has stayed

44 % in state 1, 16 % in state 2, 40 % in state 3

which is already close to the stationary probability found in (15).

If we observe the same trajectory for 100 units of time we obtain

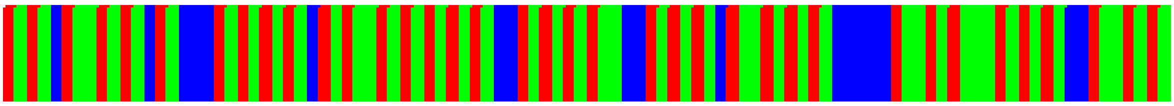

Figure 3: Trajectory after 100 units of time

that is the chain has visited

45 times state 1, 18 times state 2, 37 times state 3

that is, in percentage of time it has stayed

45 % in state 1, 18 % in state 2, 37 % in state 3

which is even closer from (15). There is actually a theoretical result saying that the transient probability (i.e. $\pi(n)$) converges geometrically fast to the stationary probability (i.e. π) as $n \to \infty$. This means that the convergence is fast *in terms* of units of time but not necessarily in terms of real time (it may be long when the unit of time is one year and very short if the unit of time is 1 nanosecond).

It should be clear that another experiment will give totally different results in Figures 2 and 3. For instance, starting from state 1 the chain may stay in state 1 for 50 consecutive units of time with the probability $(0.2)^{50}$. This is highly unlikely (!) but not impossible.