Chaînes de Markov en Terminale S

Maxime Fourny [1] &Yves Ducel [2] & Cédric Laleaux [3] & Bruno Saussereau[2]

[1] *Lycée Paul Émile Victor*
625, avenue de Gottmadingen
BP 80116
39303 CHAMPAGNOLE
maxime-simon.fourny@ac-besancon.fr

[2] *Laboratoire de Mathématiques de Besançon, Université Bourgogne Franche-Comté*
16, route de Gray
25030 Besançon cedex, FRANCE
yves.ducel@univ-fcomte.fr

[3] *Lycée Édouard Belin*
18 Rue Edouard Belin
70000 Vesoul
cedric-damien.laleaux@ac-besancon.fr

Résumé. Lorsque nous avons observé les nouveaux programmes de Terminale S en spécialité mathématiques de 2011, nous nous sommes aperçus que derrière chaque contenu probabiliste du programme se cachait une chaîne de Markov. Il n'est bien sûr pas nécessaire (ni suffisant) de présenter cette théorie générale aux élèves. Nous pensons cependant qu'il peut leur en être fait mention en fin d'année scolaire en utilisant le recul d'un an d'enseignement. Nous proposons ici une présentation de ce concept à travers quelques exemples concret dont notamment le "PageRank" de Google. Cette activité permet également, a posteriori, d'apporter une justification sur cette multiplication si "complexe" des matrices. Nous nous sommes, par la suite, également interrogés sur l'utilisation des chaînes de Markov par des outils informatiques notamment dans le cas de la matière ICN (Informatique et Création Numérique).
L'objectif de cet article est donc de présenter les chaînes de Markov à travers des exemples très concrets permettant d'en montrer l'intérêt, sans rentrer dans une théorie générale qui pour un élève de terminale S dépasserait le niveau attendu au lycée.

Mots-clés. Chaîne de Markov, PageRank, ICN...

Abstract. Markov chains have been taught in French high schools since 2011.
The purpose of this paper is to suggest a reflection about the way to teach such a theory with a very low mathematical background.
In this presentation, we first give an introduction to Markov chains with a study of Google's "PageRank" algorithm.
We then consider the applications of Markov chains to the digital creation in music or writing.
This work was done by the Probability and Statistic Reserch team of the Institut de Recherche sur l'Enseignement des Mathématiques (Mathematical Teaching Research Institute) of Franche-Comté.

Keywords. Chaîne de Markov, PageRank, ICN ...

1 Pertinence d'une Page web

Lorsque nous tapons un mot dans un moteur de recherche, nous obtenons plusieurs sites rangés par ordre de pertinence. Nous nous interrogerons ici sur cette pertinence. Comment est elle créée ?

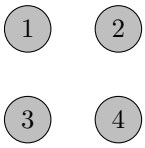
1.1 Position du problème

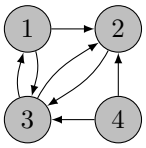
Considérons un certain sujet évoqué par quatre sites sur internet.
Discuter de la pertinence des ces pages web en rapport au sujet donné revient à leur attribuer un score : le "PageRank". Il suffit alors définir un algorithme de calcul de ce "PageRank" et d'afficher dans l'ordre les pages en rapport avec la requête.
L'idée à la base du modèle de Larry Page et Sergey Brin, fondateurs de Google, tient en deux règles :

- **R1** : Le "PageRank" doit prendre en compte les pages qui font référence dans le domaine recherché ("PageRank" élevé).

- **R2** : Une référence venant d'une page multipliant les liens doit avoir peu de crédit.

La règle **R2** nous invite à représenter les références (liens) d'une page vers une autre par une flèche, de façon à associer un graphe orienté au problème de l'attribution du "Pagerank".





En notant s_i le score de la page i, on peut considérer que le score d'une page est la somme des scores des pages pointant vers elles.
la règle **R1** est respectée. On obtient alors le système suivant à résoudre :

$$\begin{cases} s_1 = & s_3 \\ s_2 = & s_1 + s_3 + s_4 \\ s_3 = & s_1 + s_2 + s_4 \\ s_4 = & 0 \end{cases}$$

Pour respecter la règle **R2**, l'idée est considérer que plus une page envoie sur d'autres pages, moins cette référence est digne d'intérêt. On traduit cette idée en pondérant le score d'un site par l'inverse du nombre de pages pointées par ce site. Le système précédent est alors modifié de la façon suivante :

$$\begin{cases} s_1 = & \frac{s_3}{2} \\ s_2 = & \frac{s_1}{2} + \frac{s_3}{2} + \frac{s_4}{2} \\ s_3 = & \frac{s_1}{2} + s_2 + \frac{s_4}{2} \\ s_4 = & 0, \end{cases}$$

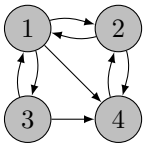
dont une solution est : $(s_1, s_2, s_3, s_4) = (\frac{2}{9}, \frac{3}{9}, \frac{4}{9}, 0)$.

Dans cette exemple, que nous appellerons **exemple 1**, le problème peut être symbolisé par la matrice A_1: $\begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \end{pmatrix}$

Définir le "PageRank" revient alors à résoudre le système :

$$A_1 \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix}$$

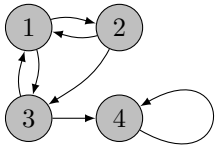
Appliquons maintenant ce principe à quatre autres exemples :
Exemple 2 :



On obtient alors les scores suivants :

$$\begin{cases} s_1 = 0.230769 \\ s_2 = 0.384615 \\ s_3 = 0.076923 \\ s_4 = 0.307692 \end{cases}$$

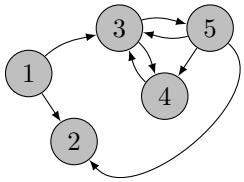
Exemple 3 :



On obtient ici :

$$\begin{cases} s_1 = 0 \\ s_2 = 0 \\ s_3 = 0 \\ s_4 = 1 \end{cases}$$

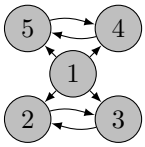
Exemple 4 :



On obtient pour cet exemple :

$$\begin{cases} s_1 = 0 \\ s_2 = 0 \\ s_3 = 0 \\ s_4 = 0 \\ s_5 = 0 \end{cases}$$

Exemple 5 :



On a $\begin{cases} s_1 = 0 \\ s_2 = 0 \\ s_3 = 0 \\ s_4 = 0.5 \\ s_5 = 0.5 \end{cases}$ Mais aussi : $\begin{cases} s_1 = 0 \\ s_2 = 0.5 \\ s_3 = 0.5 \\ s_4 = 0 \\ s_5 = 0 \end{cases}$
comme solutions.

En analysant les exemples 4 et 5 et à moindre mesure l'exemple 3, nous voyons les limites d'un tel système. Car le but de l'opération est de déterminer, dans tous les cas, une unique solution qui permette d'établir un ordre de priorité dans les scores des pages référencées.
Voilà donc comment la success story de Google aurait pu s'arrêter net dans les années 90.
Heureusement... Larry Page et Sergey Brin ont pensé à deux concepts mathématiques :

- les chaînes de Markov ;
- la formule des probabilités totales.

1.2 Chaînes de Markov

Chaînes de Markov homogène (à espace des états discret)
Intuitivement : le futur d'un processus markovien ne dépend du passé qu'à travers le présent, le passé n'a pas d'influence.

Cette idée se formalise par la donnée d'une suite de variables aléatoires réelles $(X_n)_{n\in\mathbb{N}}$ prenant leurs valeurs dans un espace E dénombrable (en général $\mathbb{N}$ ou un sous-ensemble fini de $\mathbb{N}$) qui possède la propriété suivante : pour tout entier $n \geq 0$ et pour tout $n+1$-uplet de E, $(j_0, j_1, j_2, ..., j_{n-1}, j, i) \in E^{n+1}$, on a :

$$\mathbb{P}\left(X_{n+1} = i \,|\, X_n = j, X_{n-1} = j_{n-1}, ..., X_1 = j_1, X_0 = j_0\right) = \mathbb{P}\left(X_{n+1} = i \,|\, X_n = j\right).$$

On dit qu'une telle suite de variables aléatoires réelles $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov à valeur dans E.

A priori, dans cette définition, la probabilité de passer d'un état j à un état i quelconque dépend de l'instant n. Dans la pratique du lycée, on ne considérera que des chaînes de Markov particulières (dites homogènes) telles que, pour tout i et j dans E, la probabilité de passer de l'état i à l'état j ne dépende pas de l'entier n. Dans la suite, nous ne considérerons que des chaînes de Markov homogènes.

Plusieurs utilisations des chaînes de Markov peuvent être envisagées au niveau du lycée :

- Marche aléatoire sur une échelle
- Génération aléatoire de texte
- Génération aléatoire de note de musique
- Nombreux exercices de la spécialité Mathématiques en terminale S faisant appel au calcul des probabilités.

Application de la formule des probabilités totales à une chaîne de Markov
Soit n un entier naturel fixé. Notons N le cardinal de l'espace des états E (éventuellement $N = +\infty$)
Considérons, pour tout $j \in E$, l'événement $A_j = \{X_n = j\}$.

Les événements A_j forment une partition de Ω. On a donc, pour tout $i, j \in E$,

$$\mathbb{P}(X_{n+1} = i) = \sum_{j=1}^{n} \mathbb{P}(\{X_{n+1} = i\} \cap A_j)$$

Soit :

$$\mathbb{P}(X_{n+1} = i) = \sum_{j=1}^{n} \mathbb{P}(X_{n+1} = i \,|\, X_n = j) \times \mathbb{P}(X_n = j)$$

Matrice de transition d'une chaîne de Markov homogène
Pour tout $i, j \in E$, on note $p_{i,j}$ la probabilité de passage de l'état j à l'instant n à l'état i à l'instant $n+1$:

$$p_{i,j} = \mathbb{P}\left(X_{n+1} = i | X_n = j\right)$$

On pose $P = (p_{i,j})_{i,j \in E^2}$ la matrice finie ou dénombrable (suivant que E est fini ou dénombrable) de toutes les probabilités de transition en une étape à l'instant n.
Cette matrice est appelé matrice de transition. La chaîne de Markov étant homogène, les coefficients de cette matrice ne dépendent pas de l'instant n. La matrice P est une matrice stochastique colonne.

Attention, dans la littérature (et certains exercices), on peut rencontrer la convention suivante :

$$p_{i,j} = \mathbb{P}\left(X_{n+1} = j | X_n = i\right)$$

La matrice de transition P est alors la transposée de la matrice précédente et elle est stochastique ligne.

Détermination et représentation d'une chaîne de Markov homogène
Une chaîne de Markov homogène est entièrement déterminée par la donnée de sa matrice de transition P et la loi initiale de la variable aléatoire X_0, notée π_0, considérée comme un vecteur colonne à composantes positives dont la somme est égale à 1.
Si π_n est la loi de la variable aléatoire X_n, alors on obtient grâce à la formule des probabilités totales :

$$\pi_{n+1} = P\pi_n \qquad \text{et donc :} \qquad \pi_n = P^n \pi_0$$

On notera au passage que le produit matriciel est en parfaite adéquation ici avec la formule des probabilités totales.

Avec la convention $p_{i,j} = \mathbb{P}\left(X_{n+1} = j | X_n = i\right)$, on a :

$$\pi_n^* = (\pi_0^*) P^n$$

où π_0^* est un vecteur ligne qui est le transposé du vecteur colonne π_0

Une chaîne de Markov peut être représentée par :

1. Une matrice (on vient de le voir)
2. Un graphe orienté de Markov tel que :
   - les sommets sont les différents états possibles de la chaîne ;
   - deux sommets j et i sont reliés par une flèche, avec l'étiquette p_{ij}, allant de j vers i quand la chaîne peut passer de l'état j à l'état i en une étape : i.e. : $p_{ij} > 0$.

Convergence d'une chaîne de Markov
On se rend compte assez rapidement qu'il est difficile pour une chaîne de Markov de converger presque sûrement (cas du saut de puce ou urne d'Ehrenfest).
Loi stationnaire :
Une loi de probabilité π vérifiant $\pi = P\pi$ est appelé loi stationnaire de P.
Une loi stationnaire est un vecteur propre associé à la valeur propre 1 de la matrice.

Théorème de Peyron Froebenius :

- Une chaîne de Markov est irréductible si et seulement si pour tout couple (i, j) avec $i \neq j$ de sommets du graphe, il existe un chemin de i vers j et un chemin de j vers i.
- Une chaîne de Markov irréductible à état fini possède exactement une loi de probabilité stationnaire.
- La suite de matrices $(P^k)_{k \geq 1}$ converge vers une matrice limite dont toutes les colonnes sont égales à la loi stationnaire.

Corollaire : S'il existe un entier k tel que $P^k > 0$, alors la chaîne de Markov est irréductible, et donc la suite des matrices $(P^k)_{k \geq 1}$ converge.
Ce corollaire est une condition suffisante, mais elle n'est pas nécessaire.

1.3 Applications à la pertinence d'une page web

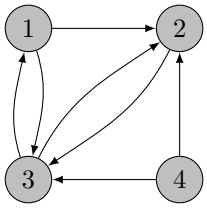
L'idée novatrice de Larry Page et Sergey Brin fut de considèrer leur modèle à partir d'un surfeur aléatoire et de se poser la question de savoir où il sera au bout de n étapes (une étape correspondant à un clic de souris sur un lien).
La suite $(X_n)_{n\in\mathbb{N}}$ est ainsi une chaîne de Markov. Il faut toutefois faire attention aux pages qui n'ont aucun lien vers l'extérieur. On remplace alors la colonne vide par un renvoi équiprobable sur une des pages de E.

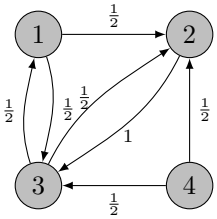
Le "PageRank" recherché n'est rien d'autre que la loi de probabilité stationnaire de la chaîne de Markov.

On remarquera que cette loi stationnaire :

- peut être déterministe (cf. exemple 3) ;
- n'existe pas forcément (cf. exemple 4) ;
- n'est pas nécessairement unique (cf. exemple 5).

Revenons de façon plus précise sur nos premiers exemples :
Retour à notre exemple 1.



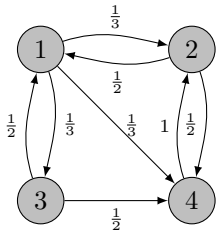


Les graphes obtenus sont des graphes de chaînes de Markov.

$$P = \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \qquad P^{100} = \begin{pmatrix} 0.222.. & 0.222.. & 0.222.. & 0.222.. \\ 0.333.. & 0.333.. & 0.333.. & 0.333.. \\ 0.444.. & 0.444.. & 0.444.. & 0.444.. \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

La loi stationnaire correspond aux scores trouvées initialement.

Retour sur l'exemple 2 :

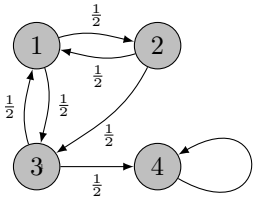


$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & 0 & 1 \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

$$P^{100} = \begin{pmatrix} 0.230769 & 0.230769 & 0.230769 & 0.230769 \\ 0.384615 & 0.384615 & 0.384615 & 0.384615 \\ 0.076923 & 0.076923 & 0.076923 & 0.076923 \\ 0.307692 & 0.307692 & 0.307692 & 0.307692 \end{pmatrix}$$

On obtient donc : $(s_1, s_2, s_3, s_4) = (0.230769, 0.384615, 0.076923, 0.307692)$.

Retour sur l'exemple 3 :

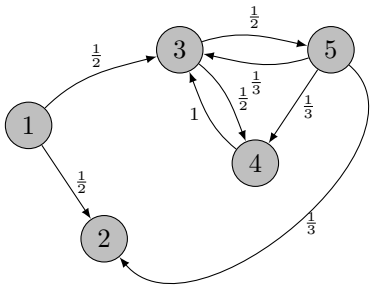


$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \end{pmatrix}$$

$$P^{100} = \begin{pmatrix} 0.000.. & 0.000.. & 0.000.. & 0.000.. \\ 0.000.. & 0.000.. & 0.000.. & 0.999.. \\ 0.000.. & 0.000.. & 0.000.. & 0.999.. \\ 0.999.. & 0.999.. & 0.999.. & 0.999.. \end{pmatrix}$$

On obtient donc : $(s_1, s_2, s_3, s_4) = (0, 0, 0, 1)$.

Retour sur l'exemple 4 :

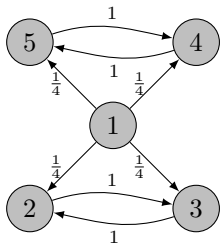


$$P = \begin{pmatrix} 0 & \frac{1}{5} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{5} & 0 & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{5} & 0 & 1 & \frac{1}{3} \\ 0 & \frac{1}{5} & \frac{1}{2} & 0 & \frac{1}{3} \\ 0 & \frac{1}{5} & \frac{1}{2} & 0 & 0 \end{pmatrix}$$

$$P^{100} = \begin{pmatrix} 0.020. & 0.020. & 0.020. & 0.020. & 0.020. \\ 0.101. & 0.101. & 0.101. & 0.101. & 0.101. \\ 0.383. & 0.383. & 0.383. & 0.282. & 0.212. \\ 0.282. & 0.282. & 0.282. & 0.282. & 0.282. \\ 0.212. & 0.212. & 0.212. & 0.212. & 0.212. \end{pmatrix}$$

Pour ce cas là on peut observer le renvoi équiprobable depuis la page 2 sur une autre page.
On obtient alors : $(s_1, s_2, s_3, s_4, s_5) = (0.020..., 0.101..., 0.383..., 0.282..., 0.212...)$.

Retour sur l'exemple 5 :



$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 1 & 0 & 0 \\ \frac{1}{4} & 1 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 & 0 & 1 \\ \frac{1}{4} & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$P^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.25 & 1 & 0 & 0 & 0 \\ 0.25 & 0 & 1 & 0 & 0 \\ 0.25 & 0 & 0 & 1 & 0 \\ 0.25 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Dans ce cinquième exemple, on s'aperçoit que $P^3 = P$ et ainsi, pour tout entier $k \geq 1$,

$$P^{2k} = P^2$$
$$P^{2k+1} = P$$

Il ne peut pas y avoir convergence de la suite des puissances de P. On notera que la chaîne de Markov considérée n'est pas irréductible.

On s'aperçoit donc que l'exemple 3 avec sa page 4 qui se comporte comme un puits et l'exemple 5 qui comporte deux portions du web ne communiquant pas entre elles rendent pour le moment ce calcul de "PageRank" inintéressant.
Et cela jusqu'à ce que les fondateurs de Google ne trouve la parade suivante :

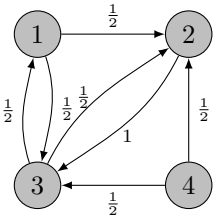
Idée Brillante de L. Page et S. Brin : Utiliser la formule des probabilités totales pour rendre la chaîne de Markov irréductible.
Leur méthode résout au passage le cas des puits car elle empêche de se retrouver dans une poche du web sans pouvoir en sortir.

Détail de l'idée :

- À chaque étape, on continue la promenade aléatoire précédente avec une probabilité de 0.85 ;
- et avec probabilité de 0.15, on fait un saut aléatoire (vers n'importe quelle page) avec une probabilité $\frac{1}{n}$ de tomber sur une page donnée, où n est le nombre de pages.

Ainsi la matrice P est remplacée par la matrice $Q = 0.85\, P + 0.15 \left(\frac{1}{n}\right)_{1 \leq i,j \leq n}$

Retour sur l'exemple 1



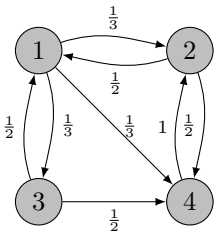
$$Q = 0.85 \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} + 0.15 \begin{pmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix}$$

$$P^{100} = \begin{pmatrix} 0.222.. & 0.222.. & 0.222.. & 0.222.. \\ 0.333.. & 0.333.. & 0.333.. & 0.333.. \\ 0.444.. & 0.444.. & 0.444.. & 0.444.. \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$Q^{100} = \begin{pmatrix} 0.208. & 0.208. & 0.208. & 0.208. \\ 0.324. & 0.324. & 0.324. & 0.324. \\ 0.401. & 0.401. & 0.401. & 0.401. \\ 0.065. & 0.065. & 0.065. & 0.065. \end{pmatrix}$$

On s'aperçoit que l'ordre de pertinence n'est pas changé par le caractère arbitraire du saut aléatoire. Toutefois le "PageRank" est légèrement modifié par rapport au modèle précédent.

Retour sur l'exemple 2

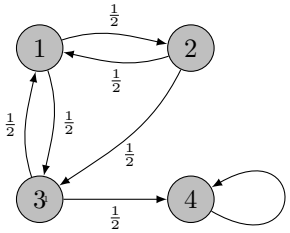


$$P^{100} = \begin{pmatrix} 0.230769 & 0.230769 & 0.230769 & 0.230769 \\ 0.384615 & 0.384615 & 0.384615 & 0.384615 \\ 0.076923 & 0.076923 & 0.076923 & 0.076923 \\ 0.307692 & 0.307692 & 0.307692 & 0.307692 \end{pmatrix}$$

$$Q^{100} = \begin{pmatrix} 0.268. & 0.268. & 0.268. & 0.268. \\ 0.412. & 0.412. & 0.412. & 0.412. \\ 0.118. & 0.118. & 0.118. & 0.118. \\ 0.344. & 0.344. & 0.344. & 0.344. \end{pmatrix}$$

$(s_1, s_2, s_3, s_4) = (0.268..., 0.412..., 0.118..., 0.344...)$.

Retour sur l'exemple 3

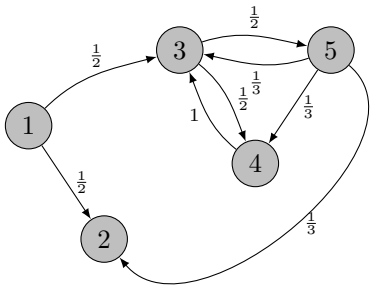


$$P^{100} = \begin{pmatrix} 0.000.. & 0.000.. & 0.000.. & 0.000.. \\ 0.000.. & 0.000.. & 0.000.. & 0.999.. \\ 0.000.. & 0.000.. & 0.000.. & 0.999.. \\ 0.999.. & 0.999.. & 0.999.. & 0.999.. \end{pmatrix}$$

$$Q^{100} = \begin{pmatrix} 0.135. & 0.135. & 0.135. & 0.135. \\ 0.095. & 0.095. & 0.095. & 0.095. \\ 0.135. & 0.135. & 0.135. & 0.135. \\ 0.633. & 0.633. & 0.633. & 0.633. \end{pmatrix}$$

$(s_1, s_2, s_3, s_4) = (0.135..., 0.095..., 0.135..., 0.633...)$. On s'aperçoit que le puits a disparu.

Retour sur l'exemple 4

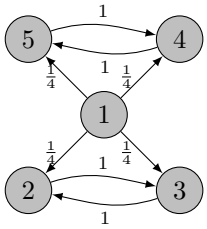


$$P^{100} = \begin{pmatrix} 0.020. & 0.020. & 0.020. & 0.020. & 0.020. \\ 0.101. & 0.101. & 0.101. & 0.101. & 0.101. \\ 0.383. & 0.383. & 0.383. & 0.282. & 0.212. \\ 0.282. & 0.282. & 0.282. & 0.282. & 0.282. \\ 0.212. & 0.212. & 0.212. & 0.212. & 0.212. \end{pmatrix}$$

$$Q^{100} = \begin{pmatrix} 0.052. & 0.052. & 0.052. & 0.052. & 0.052. \\ 0.132. & 0.132. & 0.132. & 0.132. & 0.132. \\ 0.353. & 0.353. & 0.353. & 0.353. & 0.353. \\ 0.259. & 0.259. & 0.259. & 0.259. & 0.259. \\ 0.202. & 0.202. & 0.202. & 0.202. & 0.202. \end{pmatrix}$$

$(s_1, s_2, s_3, s_4, s_5) = (0.052..., 0.132..., 0.353..., 0.259..., 0.202...)$.

Retour sur l'exemple 5



$$Q^{100} = \begin{pmatrix} 0.03 & 0.03 & 0.03 & 0.03 & 0.03 \\ 0.2425 & 0.2425 & 0.2425 & 0.2425 & 0.2425 \\ 0.2425 & 0.2425 & 0.2425 & 0.2425 & 0.2425 \\ 0.2425 & 0.2425 & 0.2425 & 0.2425 & 0.2425 \\ 0.2425 & 0.2425 & 0.2425 & 0.2425 & 0.2425 \end{pmatrix}$$

Grâce à la modification de saut aléatoire apportée sur le modèle initial, on est maintenant dans les conditions permettant d'appliquer le théorème de Peyron-Froebenius. On peut maintenant ranger les pages par pertinence. Le saut aléatoire règle également la possibilité d'être bloqué dans une portion web.
$(s_1, s_2, s_3, s_4, s_5) = (0.03..., 0.2425..., 0.2425..., 0.2425..., 0.2425...)$

2 Application des chaînes de Markov à la création numérique

2.1 Chaîne de Markov et mots

Reprenons ce concept de Markov et appliquons-le pour générer de manière automatique des mots.
Prenons un dictionnaire (français par exemple). Un tel ouvrage contient plus de 320 000 mots : tous les mots de la langue française avec les conjugaisons.

Convenons de la méthode suivante pour générer un mot par chaîne de Markov de niveau 1 :

1. Regardons (par un programme informatique) toutes les premières lettres des mots pour en extraire une distribution de fréquences. Tirons une lettre au hasard en respectant cette distribution de fréquence (ainsi la lettre "c" ou la lettre "l" aura plus de chance de sortir que la lettre "y").

   Ce sera la première lettre de notre mot.

2. Connaissant la première lettre, regardons dans le dictionnaire (toujours avec un programme informatique) la distribution des fréquences des lettres qui suivent notre première lettre. Tirons maintenant une lettre au hasard en respectant cette distribution de fréquences (qui est différente de la première distribution). La deuxième lettre dépend donc de la première.

   Et continuons comme cela pour la troisième lettre, la quatrième et ainsi de suite jusqu'à ce que nous tombions sur un caractère qui symbolise la fin d'un mot.

On peut généraliser ce principe avec des chaînes de Markov de niveau 2 : on choisit une lettre en fonction des deux précédentes; des chaînes de Markov de niveau 3 : on choisit une lettre en fonction des trois précédentes...
Voici quelques exemples de mots qu'il est possible d'obtenir en appliquant une chaîne de Markov de niveau 5 sur le dictionnaire français avec un script informatique

- aces
- ppointés
- cule
- crofilmés
- bidités
- pampreriez
- rions
- rocréontiquiez
- du
- désiras
- sous-jacentremêlassent
- énaturames

On s'aperçoit bien évidemment que la plupart des mots n'existent pas. Toutefois tous les mots ont une sonorité française. Ce sont des mots qui pourraient exister. On a numérisé l'esprit du dictionnaire français. En appliquant le même algorithme à un dictionnaire anglais, on aurait une "sonorité" des mots très british !

2.2 Chaîne de Markov et littérature

Et si nous appliquions ce principe non plus à des lettres mais à des mots ? En prenant par exemple l'œuvre "Les misérables" de Victor Hugo et en appliquant le même principe pour créer une phrase de manière aléatoire :
Connaissant un mot, un script informatique détermine les probabilités des mots qui le suivaient dans "Les misérables". Le script choisit alors un mot aléatoire en respectant cette distribution de probabilité.
Voici quelques exemples de phrases obtenues par chaîne de Markov de niveau 3 (Un mot est choisi en fonction des trois précédents) avec un script informatique.

- Tous s'y précipitèrent. Enjolras, exécutant avec sa carabine, dont il se soucie peu, laide, revêche, légitime, pleine de droits, juchée sur le code et jalouse au besoin, il n'a qu'une façon de s'en tirer et d'avoir la paix, c'est Demain.
- Et puis, chose bizarre, le premier symptôme de l'amour vrai chez un jeune homme, s'enfoncer en courant dans le crépuscule.
- Il réussit à disparaître, vendit l'argenterie de l'évêque, ne gardant que les flambeaux, comme souvenir, se glissa de ville en ville, traversa la France, vint à Montreuil-sur-Mer, eut l'idée que nous avons suivi jusqu'à ce moment, en descendit, répondit d'un air distrait aux empressements des gens de l'auberge, renvoya le cheval de monsieur est bien fatigué !

Là aussi, les phrases n'ont aucun sens, toutefois on peut admirer une certaine justesse grammaticale. De plus un spécialiste vous dira qu'on reconnaît le style d'écriture de Victor Hugo. L'esprit créatif de Victor Hugo est ainsi numérisé.
Pour aller plus loin : http://chatonsky.net/emma/ est une œuvre artistique de l'artiste Grégory Chatonsky. Un personnage féminin récite un texte qui est le fruit du croisement entre Madame Bovary et l'économiste André Orléan. Un logiciel, fondé sur des chaînes de Markov, a réalisé automatiquement ce mélange.

2.3 Chaînes de Markov et musique

On peut reprendre ce concept de chaîne de Markov avec la musique. Après avoir choisi un compositeur, on crée une partition note par note en choisissant une note par probabilité en fonction des précédentes, on obtiendrait une partition inédite qui respecterait les sonorités du compositeur choisi.
François Pachet, chercheur pour Sony, a développé un logiciel Flow Machines qui crée de la musique en analysant toutes la musicographie d'un artiste. Il a ainsi en 2016 sorti un morceau de musique : "Daddy's car" dont les spécialistes non avertis sont convaincus que c'est un inédit ... des Beatles. (source : le Figaro).