

Chapitre 11

Graphes probabilistes (Chaîne de Markov)

Sommaire

11.1 Rappels de première année	244
11.1.1 Graphes complets	247
11.1.2 Chaînes et connexité	247
11.1.3 Graphes orientés	248
11.1.4 Graphes eulériens	248
11.1.5 Comptage de chaînes	249
11.1.6 Matrice d'adjacence d'un graphe	250
11.1.7 Puissances de la matrice d'adjacence d'un graphe	251
11.1.8 Graphes et réseaux sociaux	251
11.1.9 Graphes pondérés	252
11.2 Graphes probabilistes	258
11.2.1 Un exemple	258
11.2.2 Cas général : graphes probabilistes à N états	259
11.2.3 Chaînes de Markov	260
11.2.4 Etats stables d'une chaîne de Markov homogène	263
11.3 Un cas particulier : les graphes probabilistes à 2 états	266
11.3.1 Exercices	267

- Compétences attendues. ⇨ Connaître le vocabulaire de base associé aux graphes.
⇨ Déterminer si un graphe possède une chaîne eulérienne ou un cycle eulérien.
⇨ Construire et utiliser la matrice d'adjacence d'un graphe.
⇨ Appliquer l'algorithme de Dijkstra.
⇨ Construire le graphe probabiliste et la matrice de transition associés à une chaîne de Markov.
⇨ Rechercher les états stables d'une chaîne de Markov.
⇨ Simuler une chaîne de Markov sur Python et tracer les lois empiriques et théoriques.

Le mathématicien suisse Leonhard Euler (1707-1783) montre, en 1735, l'impossibilité de passer une fois et une seule par chacun des 7 ponts de la ville de Königsberg en revenant à son point de départ. Ce problème est souvent considéré comme la naissance de la théorie des graphes mais fut longtemps considéré comme une simple curiosité mathématique.

Le développement de l'informatique et de l'algorithmique depuis les années 1950, a permis un essor très important de cette discipline auquel ont participé de nombreux mathématiciens dont Claude Berge (mathématicien français) qui a jeté les bases de la théorie des graphes par son étude sur la

théorie des jeux. On trouve des applications de cette théorie dans les réseaux informatiques, la rationalisation des services de distribution etc...

11.1 Rappels de première année

Quelques problèmes célèbres :

PROBLÈME 11.1 (Segments).

Comment tracer 5 segments sur une feuille, de telle manière que chaque segment en coupe exactement 3 autres?

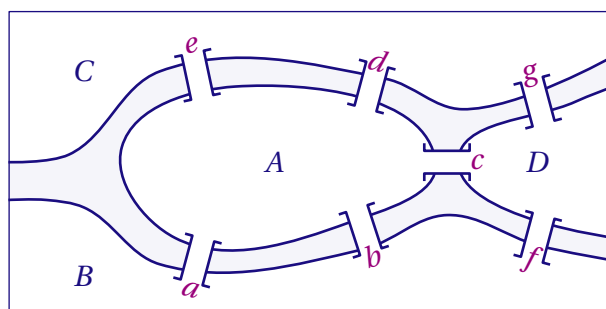
SOLUTION.

On modélise la situation à l'aide d'un graphe d'ordre 5 dont les sommets représentent les segments. Deux sommets sont adjacents si les deux segments ont une intersection.

Ainsi, il faudrait tracer un graphe d'ordre 5 dont tous les sommets seraient de degré 3. Ce qui est impossible puisque la somme des degrés de tous les sommets d'un graphe est un nombre pair.

PROBLÈME 11.2 (Ponts de Königsberg).

Au XVIII^{ème} siècle, les habitants de Königsberg (actuellement Kaliningrad, région de la Russie frontalière de la Pologne et de la Lituanie) aimaient se promener le dimanche. La ville de Königsberg comprenait 7 ponts, disposés selon le schéma ci-dessous. Le souhait des habitants de Königsberg était de faire un trajet passant une fois et une seule par chaque pont. Comment faire?



Est-il possible de se promener en ne passant qu'une seule fois sur chacun des sept ponts?

PROBLÈME 11.3 (Poignées de main).

M. et Mme EULER assistent à une réunion. Il y a trois autres couples dans l'assistance : certains participants à la réunion se saluent en se serrant la main.

- Personne ne serre sa propre main et les époux ne se serrent pas la main.
- Deux personnes quelconques de l'assemblée se serrent la main au plus une fois.
- M. EULER constate que les sept autres personnes ont échangé des poignées de mains en nombres tous distincts.

Combien de poignées de mains M. et Mme Euler ont-ils échangé avec les autres membres de la réunion?

SOLUTION.

On représente cette situation au moyen d'un graphe d'ordre 8. Ses sommets représentent les personnes présentes et une arête reliant deux sommets signifiera que les deux personnes concernées ont échangé une poignée de mains.

Comme personne ne se serre la main à lui même et les époux ne se serrent pas la main, chaque personne présente ne peut pas échanger plus de 6 poignées de mains. Le degré de chaque sommet du graphe est inférieur ou égal à 6.

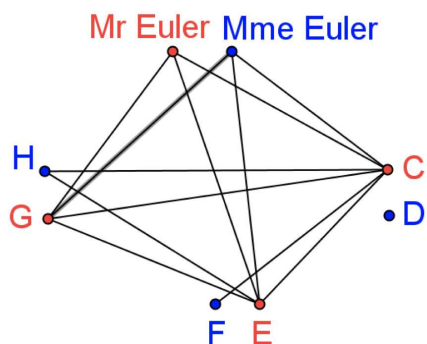
De plus, M. Euler constate que les 7 autres personnes ont échangé des poignées de mains en nombres

tous distincts; ce qui signifie que les 7 autres sommets du graphe ont des degrés tous distincts. Comme ils sont tous inférieurs ou égaux à 6; ils doivent être égaux respectivement à : 0, 1, 2, 3, 4, 5 et 6. Il y a donc une personne qui a échangé 6 poignées de mains donc son conjoint n'en a échangé aucune.

Il y a une autre personne qui a échangé 5 poignées de mains donc son conjoint en a échangé 1.

Il y a ensuite une autre personne qui a échangé 4 poignées de mains et donc son conjoint en a échangé 2.

Il reste à observer que Mme Euler a bien échangé un nombre de poignées de mains différent de 6, 0, 5, 1, 4, 2 et c'est bien exact puisque Mme Euler en a échangé 3. On remarque alors que Mr Euler en a échangé lui aussi 3.



PROBLÈME 11.4 (Problèmes d'étude de réseaux).

Un réseau informatique est un ensemble d'ordinateurs connectés entre eux. Bien sûr, tous les ordinateurs ne sont pas directement et physiquement reliés à tous les autres. Un graphe permet de décrire ces connexions. Trouver un chemin (une route en terme informatique) reliant deux ordinateurs quelconques et leur permettant d'échanger des données n'est pas une question des plus simples.

PROBLÈME 11.5 (Réseaux sociaux).

Un réseau social est un ensemble de personnes en relation entre elles. Cette relation est ici un lien de connaissance : un réseau social se représente sous la forme d'un graphe dont les sommets sont des personnes et les arêtes indiquent celles qui se connaissent. Ce graphe peut-être non-orientée (Facebook) ou au contraire orientée (Twitter : la relation est « être abonné au fil d'un autre compte »). Dans un réseau social, tous les individus n'ont pas les mêmes positions, certains ont plus d'influence que d'autres. Pour quantifier cette influence, c'est-à-dire pour rechercher des influenceurs, on introduit les notions de centralité et d'intermédiarité. Il est difficile d'imaginer le graphe complet indiquant les relations entre quelques 7 milliards de sommets représentant chacun une personne. Stanley Milgram a proposé et réalisé une expérience montrant le concept des « six degrés de séparation », théorie établie initialement par le hongrois Karinythy en 1929, idée qui est restée et est maintenant connue sous le nom de "paradoxe de Milgram" ou encore de « effet du petit monde » (ou « small world effect »).

Remarques.

De très nombreux problèmes pratiques peuvent être ainsi schématisés à l'aide d'un graphe; en simplifiant la représentation, on peut ainsi trouver plus rapidement la solution (ou en voir l'impossibilité!). Pour certains problèmes, comme ceux que nous venons de voir, les arêtes n'ont pas d'orientation. Pour d'autres, il est indispensable d'avoir une orientation sur le graphe : le plan d'une ville comme graphe non orienté satisfera le piéton, tandis que ce même graphe orienté par les sens de circulation sera bien plus apprécié de l'automobiliste.

Une question importante est celle du choix du graphe associé à une situation donnée (il peut y en avoir plusieurs); comment choisir les sommets et les arêtes? Comme on vient de le voir dans le

problème des segments, ce n'est pas toujours évident. Dans des paragraphes ultérieurs, on étudiera des questions de compatibilité, il faudra décider si les arêtes correspondent aux couples de points compatibles ou incompatibles, et si les arêtes sont orientées ou non.

Nous allons formaliser les notions qui précèdent.

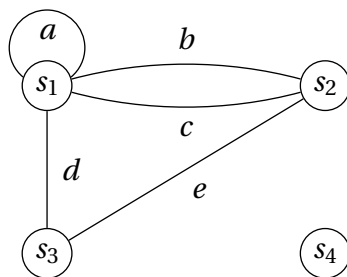
Définition 11.1 (Graphe, sommets, arêtes, sommets adjacents).

Un graphe G (non orienté) est constitué d'un ensemble S de points appelés *sommets*, et d'un ensemble A d'*arêtes*, tels qu'à chaque arête sont associés deux sommets, appelés ses *extrémités*. Deux sommets qui sont les extrémités d'une arête sont dits *adjacents*.

Les deux extrémités peuvent être distinctes ou confondues; dans ce dernier cas, l'arête s'appelle une *boucle*. Deux arêtes peuvent aussi avoir les mêmes extrémités (on dit alors qu'elles sont *parallèles*).

Exemple 11.1. On considère le graphe G_1 , de la figure 11.7. Le sommet s_4 est un sommet isolé, l'arête a est une boucle, b et c sont des arêtes ayant mêmes extrémités, les sommets s_1 et s_2 sont adjacents, ainsi que s_1 et s_3 , puisqu'ils sont reliés par une arête.

FIGURE 11.1 – Le graphe G_1



Remarques.

La position des sommets et la longueur ou l'allure des arêtes n'ont aucune importance.

Déterminer les graphes identiques est un problème très difficile en général, dès que le nombre de sommets est assez grand.

Une première manière d'évaluer la complication d'un graphe est de compter le nombre de ses sommets :

Définition 11.2 (Ordre d'un graphe).

L'*ordre* d'un graphe est le nombre de ses sommets.

Définition 11.3 (Degré d'un sommet, parité d'un sommet).

On appelle *degré d'un sommet* le nombre d'arêtes dont ce sommet est une extrémité (les boucles étant comptées deux fois). Un sommet est *pair* (respectivement *impair*) si son degré est un nombre pair (respectivement impair).

Exemple 11.2. Dans la figure 11.7, s_1 est de degré 5, s_2 de degré 3, s_4 de degré 0.

On prouve facilement le théorème suivant :

Théorème 11.1 (Formule d'Euler ou formule des poignets de main).

Soit n et p deux entiers naturels non nuls.

On considère un graphe constitué de n sommets notés $A_1, \dots, A_n$ et de p arêtes.

En désignant, pour tout $i \in \llbracket 1, n \rrbracket$, par d_i le degré du sommet A_i , alors on a : $\sum_{i=1}^n d_i = 2p$.

La somme des degrés de tous les sommets d'un graphe est égale à deux fois le nombre d'arêtes de ce graphe; c'est donc un nombre pair.

Preuve. Lorsque on additionne les degrés des sommets, chaque arête est comptée deux fois, une fois pour chaque extrémité. $\diamond$

Proposition 11.2. Dans un graphe le nombre de sommets impairs est toujours pair.

Preuve. En effet, sinon, la somme des degrés des sommets serait impaire. $\diamond$

11.1.1 Graphes complets

Définition 11.4 (Graphe complet).

Un graphe (simple) est dit *complet* si tous ses sommets sont *adjacents*, c'est-à-dire si toutes les arêtes possibles existent. On appellera K_n le graphe complet à n sommet (il n'y en a qu'un).

11.1.2 Chaînes et connexité

Dans bien des problèmes de graphes, il est naturel de considérer ce que l'on peut appeler, de façon informelle, des « parcours » ou « chemins ». Le mot utilisé en théorie des graphes est *chaîne*. La notion intuitive de chaîne, ou plus tard de chaîne orientée, se comprend bien sur un dessin.

Définition 11.5 (Chaîne, longueur d'une chaîne, cycle).

Une *chaîne* dans un graphe G est une suite finie : $s_0; a_1; s_1; a_2; s_2; a_3; s_3; \dots; a_n; s_n$ débutant et finissant par un sommet, alternant sommets et arêtes de telle manière que chaque arête soit encadrée par ses sommets extrémités.

La *longueur* de la chaîne est égale au nombre d'arêtes qui la constituent; la chaîne est *fermée* si $s_0 = s_n$, si de plus toutes ses arêtes sont distinctes on dit alors que c'est un *cycle*.

Remarque. Quand il n'y a pas d'ambiguïté (pas d'arêtes multiples), on peut définir une chaîne par seulement la suite de ses sommets ou par seulement la suite de ses arêtes.

Enfin on a :

Définition 11.6 (Graphe connexe). Un graphe est *connexe* si deux sommets quelconques sont reliés par une chaîne.

Définition 11.7 (Distance entre deux sommets). Soit G un graphe connexe, s et s' , deux sommets quelconques de G . Le graphe étant connexe, il existe au moins une chaîne reliant s et s' . On appelle *distance entre s et s'* la plus petite des longueurs des chaînes reliant s à s' .

Remarque. Lorsque le graphe n'est pas connexe, il existe au moins deux sommets qui ne sont pas reliés par une chaîne. On dit parfois que la distance entre ces sommets est infinie.

Définition 11.8 (Diamètre d'un graphe). On appelle *diamètre* d'un graphe connexe, la plus grande distance entre ses sommets.

Remarque. Lorsque le graphe n'est pas connexe, on dit parfois que son diamètre est infini.

11.1.3 Graphes orientés

Il existe de nombreux domaines où les graphes sont orientés. Par exemple : plan de ville, avec les sens interdits ; parcours en montagne, où il est utile d'indiquer le sens de montée ! Circuit électrique en courant continu, où il faut orienter les arêtes pour décider du signe de l'intensité : ce n'est pas la même chose de faire passer 10 ampères de A vers B ou de B vers A ; graphe d'ordonnancement, où les arêtes relient une tâche à une autre qui doit la suivre : on ne peut faire la peinture avant le plâtre.

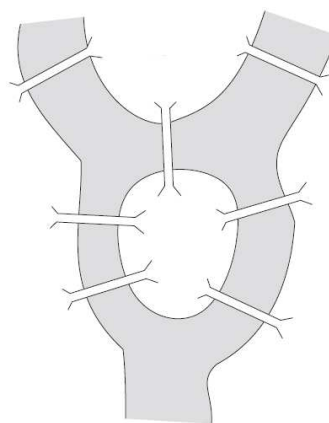
Définition 11.9. On appelle graphe orienté un graphe où chaque arête est orientée, c'est-à-dire qu'elle va de l'une des ses extrémités, appelée *origine* ou *extrémité initiale* à l'autre, appelée *extrémité terminale*.

Dans un graphe orienté, chaque arête orientée possède un début et une fin. Toutes les notions que nous avons définies pour un graphe ont un équivalent pour un graphe orienté. Nous nous contenterons de rajouter le mot « orienté » pour préciser ; le contexte rendra évidente l'interprétation à donner.

En particulier, une chaîne orientée est une suite d'arêtes telle que l'extrémité finale de chacune soit l'extrémité initiale de la suivante. On prendra garde au fait que l'on peut définir et utiliser des chaînes (non orientées) sur un graphe orienté. Par exemple, sur un plan de ville où toutes les rues sont en sens unique, un parcours de voiture correspond à une chaîne orientée, un parcours de piéton correspond à une chaîne (non orientée).

11.1.4 Graphes eulériens

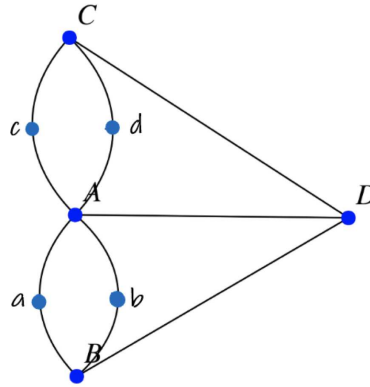
PROBLÈME 11.6 (Les sept ponts de Königsberg). Le problème qui suit est, selon la légende¹, à l'origine de l'invention des graphes par EULER, qui résidait à Königsberg. Au XVIII^e siècle, les habitants de Königsberg (actuellement Kaliningrad, région de la Russie frontalière de la Pologne et de la Lituanie) aimaient se promener le dimanche. La ville de Königsberg comprenait sept ponts, disposés selon le schéma de la figure ci-contre. Le souhait des habitants de Königsberg était de faire un trajet passant une fois et une seule par chaque pont. Comment faire ?



SOLUTION.

On considère les 4 zones A, B, C, D du plan, on schématise le problème par un graphe à 4 sommets A, B, C, D . Ces sommets sont tous de degré impair (A de degré 5, B, C, D de degré 3 chacun.) Il

1. EULER a bien trouvé la solution générale de ce type de problèmes mais sans utiliser les graphes et, s'il a prouvé dans quels cas un tel trajet était impossible, il n'a pas prouvé pourquoi dans les autres cas c'est toujours possible.



n'existe donc pas de chemin eulérien.

Définition 11.10 (Chaîne eulérienne). Une chaîne est *eulérienne* si elle contient une fois et une seule chaque arête du graphe; si la chaîne est un cycle (sommet de départ et de fin confondus), on l'appelle *cycle eulérien*.

Le théorème suivant, dit théorème d'EULER, qu'on admettra, est à l'origine de la théorie des graphes :

Théorème 11.3 (d'EULER). *Un graphe connexe a une chaîne eulérienne si et seulement si tous ses sommets sont pairs sauf au plus deux.*

De façon plus précise :

- si le graphe n'a pas de sommet impair, alors il a un cycle eulérien;
- le graphe ne peut avoir un seul sommet impair;
- si le graphe a deux sommets impairs, ce sont les extrémités de la chaîne eulérienne.

La propriété suivante, conséquence immédiate du théorème, est souvent utile :

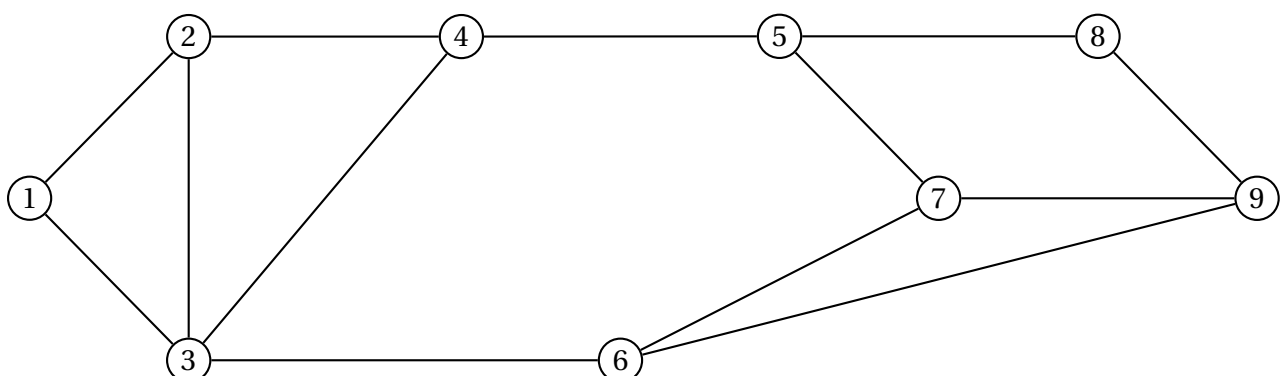
Proposition 11.4. *Un graphe connexe ayant plus de deux sommets impairs ne possède pas de chaîne eulérienne.*

Ces résultats permettent de résoudre beaucoup de problèmes pratiques se traitant en théorie des graphes.

11.1.5 Comptage de chaînes

Un problème

Sébastien se rend régulièrement en train de la ville U (sommet 1) à la ville V (sommet 9), dans un réseau donné par le graphe ci-dessous. Il fait toujours le trajet en 5 étapes, et veut faire à chaque fois un chemin différent. Combien de trajets pourra-t-il faire ?



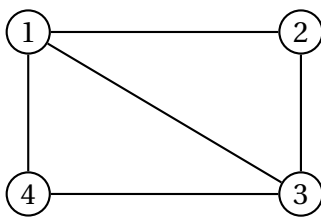
Une solution

Si l'on essaie une énumération complète de tous les chemins de longueur 5 de U à V , on se rend vite compte qu'il est difficile de ne pas en oublier. Il faut trouver un moyen systématique de faire l'énumération. La solution passe par les matrices.

11.1.6 Matrice d'adjacence d'un graphe

Définition 11.11. Soit G un graphe qui possède n sommets numérotés de 1 à n . On appelle matrice d'adjacence du graphe la matrice $A = (a_{i;j})$, où $a_{i;j}$ est le nombre d'arêtes joignant le sommet de numéro i au sommet de numéro j .

Exemple 11.3. Le graphe G ci-dessous a la matrice d'adjacence A ci-contre.



$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Remarque. L'allure de la matrice d'adjacence donne des indications sur la nature du graphe :

- La matrice d'adjacence d'un graphe sans boucle n'a que des 0 sur la diagonale.
- La matrice d'adjacence d'un graphe sans arête parallèle n'a que des 1 ou des 0.
- La matrice d'adjacence d'un graphe non orienté est symétrique par rapport à sa diagonale.
- La matrice d'adjacence d'un graphe complet n'a que des 1, hormis sur sa diagonale où il y a des 0.

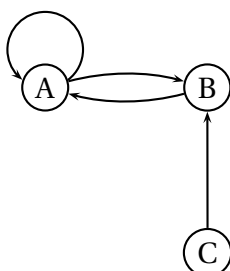
Cas d'un graphe orienté On a vu que la matrice d'adjacence d'un graphe non orienté est symétrique. Il n'en va pas de même pour celle d'un graphe orienté car si une arête a pour extrémité initiale le sommet A et pour extrémité finale le sommet B , il n'en existe pas forcément une allant de B vers A , aussi dans ce cas le coefficient correspondant aux arêtes allant de A vers B ne sera pas le même que le coefficient correspondant aux arêtes allant de B vers A .

On donc doit convenir d'un sens de lecture pour la matrice d'un graphe orienté. La matrice

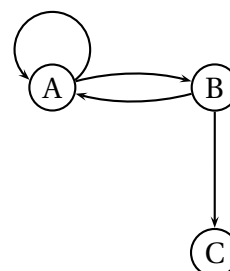
$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

peut correspondre à l'un ou l'autre des graphes ci-dessous selon qu'on choisit de lire les extrémités initiales en ligne ou en colonne :

En ligne



En colonne



On convient alors de la chose suivante :

Définition 11.12. Par convention, dans la matrice d'adjacence d'un graphe G orienté, le terme a_{ij} (i -ième ligne, j -ième colonne) désigne le nombre d'arêtes d'origine le sommet i et d'extrémité finale le sommet j .

11.1.7 Puissances de la matrice d'adjacence d'un graphe

Propriété 11.5. Soit G un graphe de matrice d'adjacence A . Le nombre de chaînes de longueur n joignant le sommet i au sommet j est donné par le terme d'indice $i; j$ de la matrice A^n .

SOLUTION.

On trouve la matrice d'adjacence A et on calcule A^5 , le coefficient $a_{1,9} = 15$ ce qui donnerait 15 chemins possibles.

Théorème 11.6 (Condition nécessaire et suffisante de connexité).

Si G est un graphe orienté dont les sommets sont les éléments de $\llbracket 1, n \rrbracket$ et dont la matrice d'adjacence est notée M , alors G est connexe si et seulement si la matrice $I_n + M + \dots + M^{n-1}$ a tous ses coefficients strictement positifs.

11.1.8 Graphes et réseaux sociaux

Considérons un réseau social, pour lequel les membres peuvent être représentés par des sommets et les liens d'amitié entre eux par des arêtes, le graphe associé est alors non orienté car le lien d'amitié n'est validé que par consentement mutuel.

Considérons un autre réseau social, style Twitter, où un membre peut suivre l'actualité d'un autre sans qu'il y ait réciprocité, alors le graphe associé est cette fois-ci orienté.

Dans un réseau social, tous les individus n'ont pas les mêmes, positions, certains ont plus d'influence que d'autres. Pour rechercher des influenceurs, on introduit les notions de centralité et d'intermédiarité.

Définition 11.13 (Centralité).

Soit G un graphe non orienté, constitué de n sommets notés, $s_1, s_2, \dots, s_n$, ($n \geq 2$) et on note, pour tout $i \in \llbracket 1, n \rrbracket$, d_i le degré du sommet s_i .

On appelle degré de centralité du sommet s_i , le réel $\frac{d_i}{n-1}$.

Remarques. Le degré de centralité du sommet s_i est donc égal au nombre de sommets adjacents à s_i divisé par le nombre de sommets du graphe autres que s_i . Cette quantité est donc comprise entre 0 et 1.

Un degré proche de 1 signifie une forte centralité.

La mesure du degré renseigne sur l'importance du sommet. Un sommet ayant un degré élevé est considéré comme central car relié à beaucoup d'autres sommets. Il a un fort potentiel d'attraction. Si le graphe est orienté, chaque sommet possède deux degrés de centralité, l'un comptant les liens entrants et l'autre comptant les liens sortants.

Définition 11.14 (Intermédiarité).

Soit G un graphe, constitué de n sommets notés, $s_1, s_2, \dots, s_n$, ($n \geq 2$) et soit $k \in \llbracket 1, n \rrbracket$, notons pour tout i et j de $\llbracket 1, n \rrbracket$, $n_{i,j}$ le nombre total de plus courts chemins entre s_i et s_j et $n_{i,j}(s_k)$ le nombre de plus courts chemins entre s_i et s_j passant par s_k .

On appelle degré d'intermédiarité du sommet s_k , le réel
$$\sum_{i=1, i \neq k}^n \sum_{j=1, j \neq k}^n \frac{n_{i,j}(s_k)}{n_{i,j}}.$$

Remarque. Le degré d'intermédiarité d'un sommet est donc la somme des fréquences d'apparition du sommet dans l'ensemble des plus courts chemins entre deux sommets quelconques du graphe.

11.1.9 Graphes pondérés

Définitions

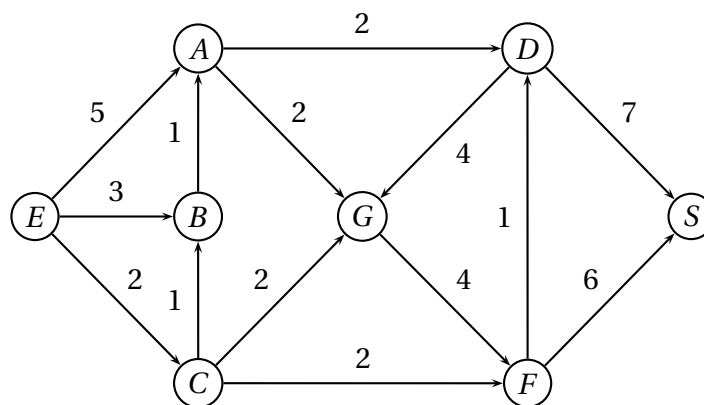
Définition 11.15. On appelle *graphe pondéré* un graphe tel que, à chaque arête a est associé un poids P_a .

Remarque. En ECG nous nous limiterons aux graphes pondérés par des poids positifs.

Les applications des graphes pondérés sont nombreuses : cartes routières avec des indications de durée, de tarif ou de distance portée sur des routes entre deux lieux, par exemple.

PROBLÈME 11.7.

Le graphe suivant représente un réseau routier (avec des sens interdits) ; quel est l'itinéraire le plus court qui relie E à S ?



Remarque. Attention à la terminologie : Ainsi, selon la terminologie exacte, la chaîne $EAGF$ est une chaîne de *longueur* 3 (trois arêtes) mais de *poids* 11 (somme des poids des arêtes).

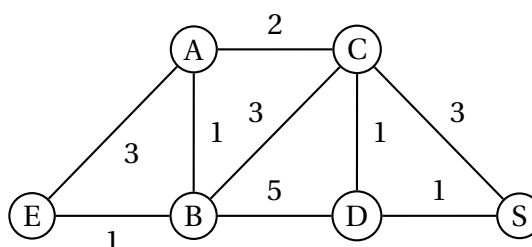
L'algorithme de DIJKSTRA

Ce problème est très naturel, et se présente dans bien des domaines différents. Même pour un graphe aussi simple que celui-là, il n'est pas évident d'être sûr que l'on a trouvé le plus court chemin. La méthode la plus immédiate est de considérer tous les chemins de A à S , et de chercher le plus court. Cette méthode est très inefficace : on peut bien sûr se limiter aux chemins sans cycle, donc de longueur bornée par le nombre de sommets, mais, même comme cela, le nombre de chemins possibles reste très grand. Voir la figure 11.2 page ci-contre représentant l'arbre pondéré de tous les chemins sans cycle menant de E à S .

Il nous faut donc obtenir une méthode plus efficace qui, si possible, donne systématiquement le bon résultat.

L'une d'elles est l'algorithme de DIJKSTRA.

Exemple



La figure 11.3 page 254 présente tous les chemins sans cycle allant de E à S . On pourra y suivre le fonctionnement de l'algorithme.

FIGURE 11.2 – Chemins sans cycle menant de E à S du problème 11.7

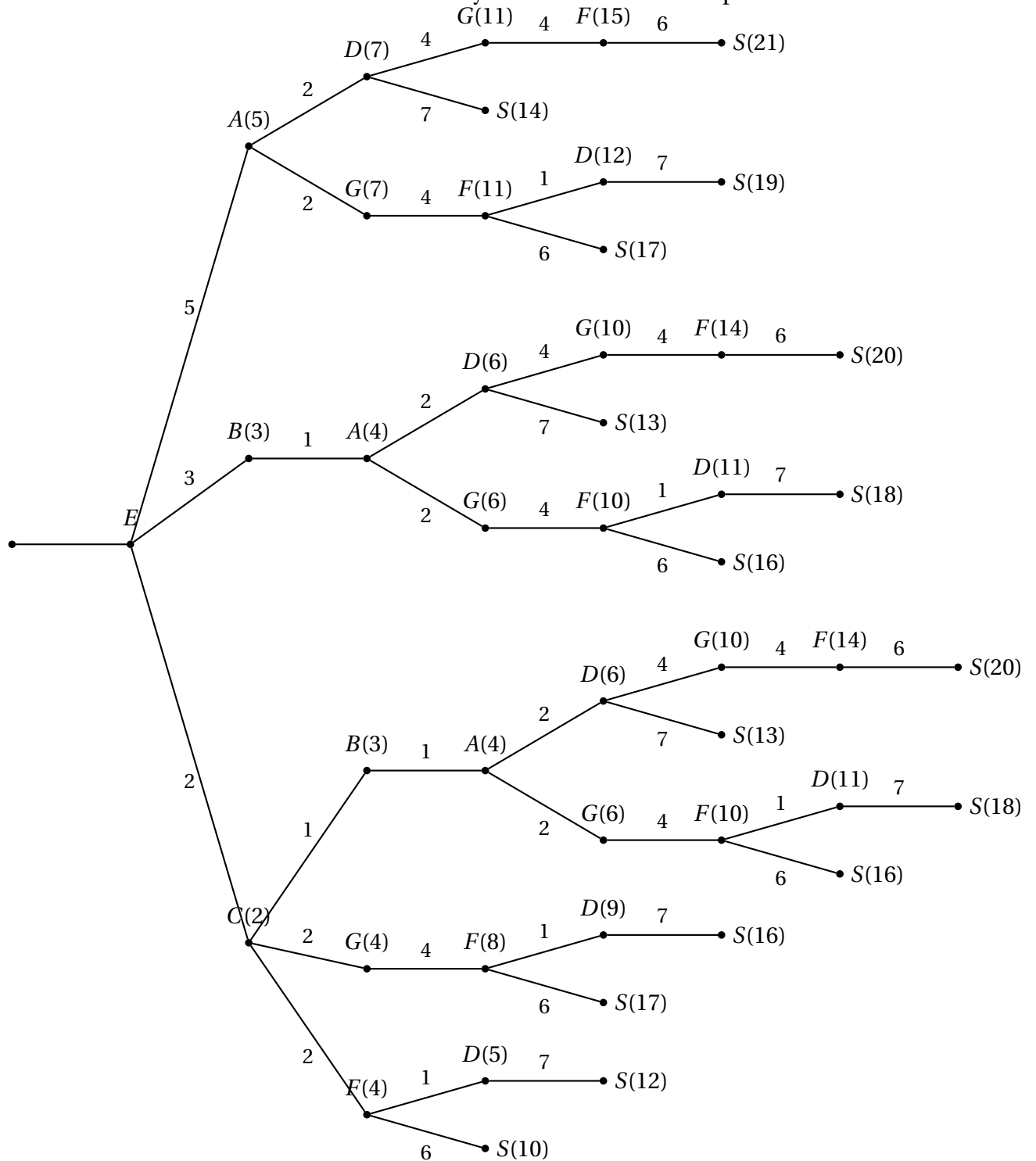
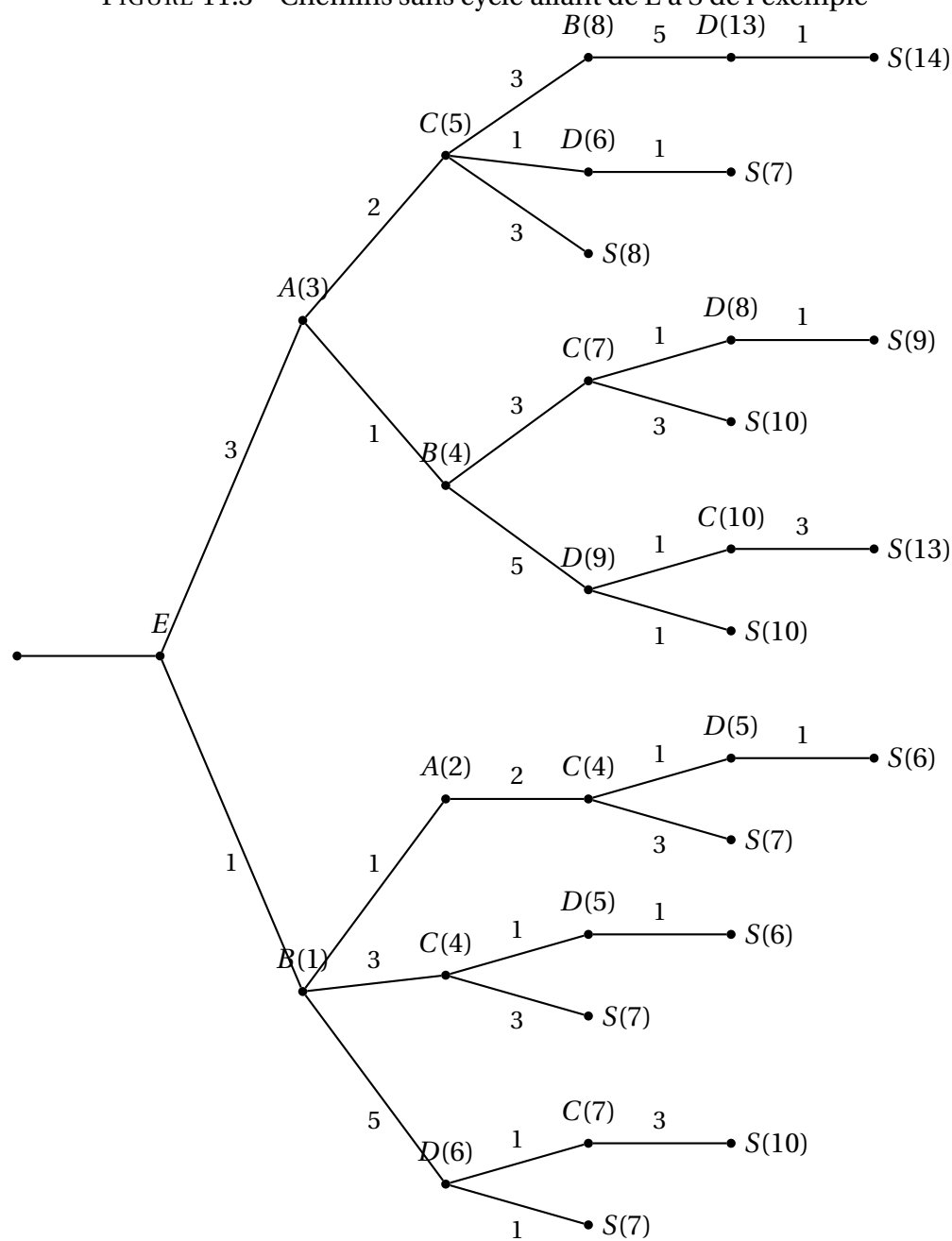


FIGURE 11.3 – Chemins sans cycle allant de E à S de l'exemple



Initialisation : On affecte *provisoirement* le poids 0 à E et on attribue *provisoirement* à tous les sommets le poids $+\infty$.

Algorithme : Tant que tous les sommets ne sont pas marqués définitivement ou que le sommet S à atteindre n'est pas affecté du plus petit des poids provisoires, exécuter les actions suivantes :

- parmi tous les sommets provisoirement pondérés, fixer **définitivement** le poids d'un de ceux qui ont un poids minimum ; soit T ce sommet.
- parmi tous les sommets T' *non marqués définitivement* adjacents à T , si la somme s de T et de l'arête de T à T' est inférieure à leur poids provisoire les affecter de s (en indiquant entre parenthèse la provenance)

Le tableau 11.1 page suivante montre comment appliquer cet algorithme à notre exemple.

Pour retrouver le chemin le plus court, on se sert des sommets marqués : Sur S l'indication est qu'on provient de D , sur D qu'on provient de C , sur C qu'on provient de B et sur B qu'on provient de E .

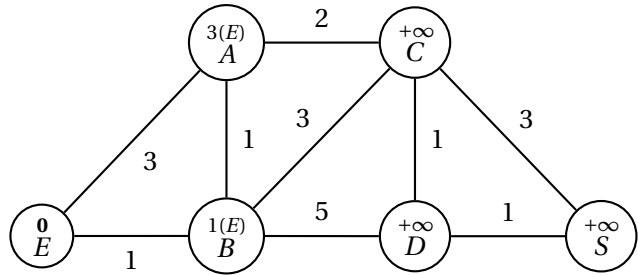
La plus courte chaîne est donc $EBCDS$, elle mesure 6.

Remarque. On notera que l'algorithme fourni aussi la plus courte chaîne pour aller de E à n'importe lequel des sommets du graphe.

TABLE 11.1 – Application de l'algorithme de DIJKSTRA

Étape 1 Aucun sommet n'est marqué définitivement. E est le seul marqué provisoirement. On marque donc définitivement son poids (0) et on affecte provisoirement A et B , adjacents à E de la somme de E (0) et du poids des arêtes menant de E à A et de E à B .

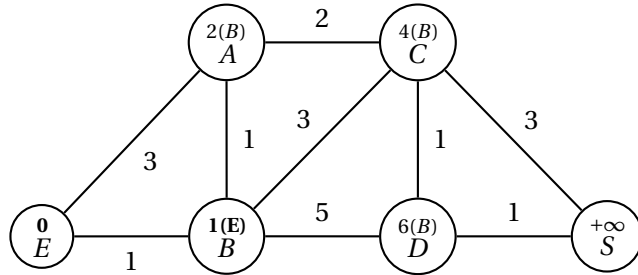
On obtient le graphe marqué ci-contre.



Étape 2 B est le plus petit poids parmi les poids provisoires. On marque donc définitivement son poids (1). On affecte les sommets adjacents à B des poids suivants :

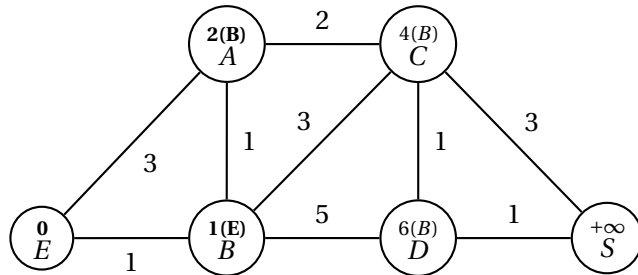
- $A = 2$ car $1(B) + 1(\text{arête } BA) < 3$ (poids provisoire de A)
- $C = 4$ ($1(B) + 3(BC)$)
- $D = 6$ ($1(B) + 5(BD)$)

On obtient le graphe marqué ci-contre.



Étape 3 A est le plus petit poids parmi les poids provisoires. On marque donc définitivement son poids (2). On ne change rien à C $2(A) + 2(AC) = 4$ (on pourrait aussi le changer car il y a égalité)

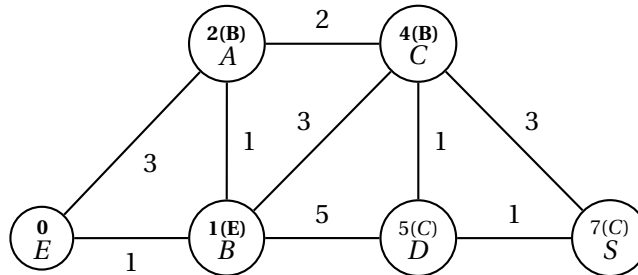
On obtient le graphe marqué ci-contre.



Étape 4 C est le plus petit poids parmi les poids provisoires. On marque donc définitivement son poids (4). On affecte

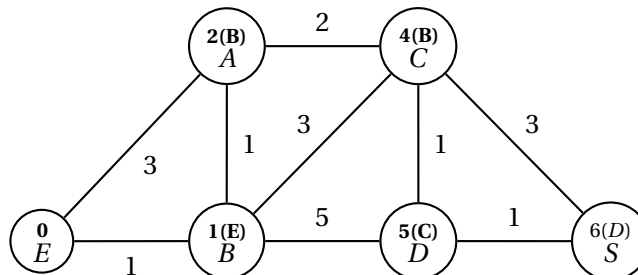
- D du poids provisoire 5 car $4(C) + 1(CD) < 6$
- S du poids provisoire 7 ($4(C) + 3(CS)$)

On obtient le graphe marqué ci-contre.



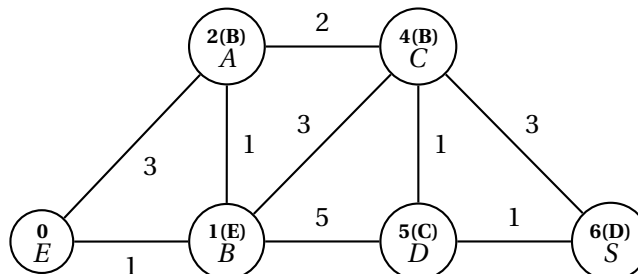
Étape 5 D est le plus petit poids parmi les poids provisoires. On marque donc définitivement son poids (5). On affecte S du poids provisoire 6 car $5(D) + 1(DS) < 7$

On obtient le graphe marqué ci-contre.



Étape 6 S est le plus petit poids parmi les poids provisoires. On marque donc définitivement son poids (6).

On obtient le graphe marqué ci-contre.



L'algorithme se présente aussi sous forme de tableau :

E	A	B	C	D	S
0	3(E)	1(E)	∞	∞	∞
	2(B)	1 (E)	4(B)	6(B)	
	2 (B)				
			4 (B)	5(C)	7(C)
				5 (C)	6(D)
					6 (D)

Théorème 11.7. Soit G un graphe connexe, éventuellement orienté, l'algorithme de DIJKSTRA fournit systématiquement la plus courte chaîne d'un sommet initial à tous les autres sommets du graphe.

On l'admettra.

Exemple 11.4.

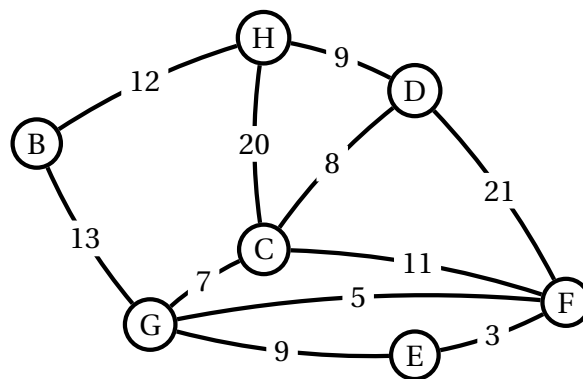
Des touristes sont logés dans un hôtel H .

Un guide souhaite leur faire visiter la région en empruntant les routes signalées comme d'intérêt touristique signalées par l'Office du tourisme.

Les tronçons de route qu'il souhaite emprunter sont représentées sur le graphe ci-contre. Sur chaque arête figure la longueur en km des différents tronçons.

Un musée est situé en E .

Déterminer le plus court chemin menant de l'hôtel H au musée E .



H	B	D	F	E	G	C	choix du sommet	coefficient
0(H)	∞	∞	∞	∞	∞	∞	H	0
X	12(H)	9(H)	∞	∞	∞	20	D	9
X	12(H)	X	30(D)	∞	∞	17(D)	B	12
X	X	X	30(D)	∞	25(B)	17(D)	C	17
X	X	X	28(C)	∞	24(C)	X	G	24
X	X	X	28(C)	33(G)	X	X	F	28
X	X	X	X	31(F)	X	X	E	31

Le plus court chemin pour aller de H vers E mesure 31 km : c'est H - D - C - F - E .

On est effectivement arrivé à E en venant de F , à F en venant de C , à C en venant de D et à D en venant de H .

11.2 Graphes probabilistes

11.2.1 Un exemple

Le problème : évolution de population

Deux villes X et Y totalisent à elles deux une population d'un million d'habitants.

La ville X est plus agréable, mais la ville Y offre de meilleurs salaires.

20 % des habitants de Y partent chaque année habiter X pour avoir un meilleur cadre de vie, et 5 % des habitants de X partent chaque année habiter Y pour augmenter leur niveau de vie.

À l'année 0, un quart des habitants sont en X .

On se pose les questions suivantes :

- Comment sera répartie la population, entre les villes X et Y au bout de 1, 2, 5, 10, 30, 40 ans?
- Que se serait-il passé au bout de 1, 2, 5, etc. ans si 99 % des habitants avaient été initialement en X (ou en Y) ?
- Que se serait-il passé au bout de 1, 2, 5, etc. ans si la population avait été également répartie entre les deux villes en l'année zéro ?

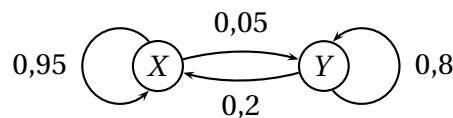
Une méthode de résolution

1. L'énoncé nous dit que 95 % des gens qui sont en X y restent, 5 % partent en Y , et que 80 % des gens qui sont en Y y restent, 20 % partent en X .

En appelant X_n la population de la ville X à l'année n et Y_n celle de Y , on peut représenter l'évolution par le système d'équations :

$$\begin{cases} X_{n+1} = 0,95X_n + 0,2Y_n \\ Y_{n+1} = 0,05X_n + 0,8Y_n \end{cases}$$

2. On peut représenter la situation par le graphe suivant, où l'on a marqué, sur chaque arête joignant le sommet X au sommet Y , la proportion de population qui passe à chaque étape de X à Y . Remarquons que, puisque la population ne peut disparaître ou apparaître, la somme des coefficients sur toutes les arêtes quittant un sommet doit être 1 :



Si l'on note, pour tout entier naturel n , $P_n = (X_n \ Y_n)$ le vecteur ligne qui décrit la population de X et de Y au bout de n années, déterminer la matrice M telle que le système d'équation d'évolution trouvée précédemment peut se réécrire :

$$P_{n+1} = P_n \times M$$

On appellera M : *matrice de transition du système*.

Attention! Le produit ne se fait pas à droite de la matrice, comme on en a l'habitude, mais à gauche. Cela présente l'avantage de garder l'écriture des vecteurs en ligne, et c'est l'habitude en probabilité.

On trouve $M = \begin{pmatrix} 0,95 & 0,05 \\ 0,2 & 0,8 \end{pmatrix}$.

- (a) Exprimons P_n en fonction de P_0 , de M et de n .

On montre facilement par récurrence que $P_n = P_0 \times M^n$.

- (b) À l'aide de cette formule, on peut alors répondre aux questions qu'on se pose dans le problème.
On a $P_1 = P_0 M$, $P_2 = P_0 M^2$, $P_5 = P_0 M^5$, $P_{10} = P_0 M^{10}$, $P_{30} = P_0 M^{30}$, $P_{40} = P_0 M^{40}$ avec $P_0 = (250\,000\,750\,000)$ ou $P_0 = (990\,000\,10\,000)$, ou $P_0 = (10\,000\,990\,000)$, ou $P_0 = (500\,000\,500\,000)$.

3. (a) On peut constater que, quand n devient grand, quelle que soit la répartition de population à l'année 0 le système converge vers $(800\,000\ 200\,000)$.
- (b) On suppose que $P_0 = (800\,000\ 200\,000)$.
On a alors $P_1 = P_2 = P_3 = \dots = P_n = P_0$.
Cette répartition ou cet état est stable.
- (c) On a vu que, quelle que soit la population de départ, le système converge vers cet état. Il peut donc être intéressant d'être en mesure de le déterminer dès le départ.
On admet qu'une telle répartition existe et on appelle x la population de X et y celle de Y pour lesquelles la population de chaque ville est la même chaque année.
- i. Que devient le système d'équations d'évolution si $(X_n\ Y_n) = (x\ y)$? On obtient comme système :

$$\begin{cases} x &= 0,95x + 0,2y \\ y &= 0,05x + 0,8y \end{cases}$$

ii. Alors x et y sont solutions de : $\begin{cases} 0,2y &= 0,05x \\ x + y &= 1\,000\,000 \end{cases}$

- iii. Ce qui donne bien $x = 800\,000$ et $y = 200\,000$.

11.2.2 Cas général : graphes probabilistes à N états

On considère un système qui peut se trouver dans N états $\{1; 2; \dots; N\}$, avec une certaine probabilité, variable au cours du temps, pour chaque état.

On s'intéresse à l'évolution de ce système au cours du temps, et on fait l'hypothèse que la probabilité de transition de l'état i à l'état j est indépendante du temps, et ne dépend pas de l'histoire antérieure, mais seulement de l'état dans lequel on se trouve.

De bons exemples de tels systèmes sont donnés par les jeux de hasard, tels que jeu de l'oie, Monopoly, jacquet, petits chevaux, etc. Pour de tels jeux, l'état est donné par la case sur laquelle on se trouve; la façon dont on y est arrivé n'a pas d'importance pour la suite du jeu.

Les systèmes qu'on observe dans la vie réelle sont en général beaucoup plus complexes, mais l'approximation simple qu'on en fait ici donne souvent des indications utiles; ce type de modèle est utilisé en pratique dans un grand nombre de situations, avec de bons résultats.

On peut représenter un tel système par un graphe orienté, dont les sommets sont les états du système, et où l'on associe à chaque transition, de l'état i à l'état j , une arête orientée allant de i vers j , étiquetée par la probabilité de transition, c'est-à-dire la probabilité conditionnelle d'être dans l'état j à l'instant $n + 1$ sachant que l'on est dans l'état i à l'instant n . Remarquons que l'on peut rester dans un même état : le graphe peut avoir des boucles.

Définition 11.16 (Graphe probabiliste).

On appelle *graphe probabiliste* un graphe orienté, tel que pour chaque couple de sommets $(i; j)$ distincts ou confondus il existe au plus une arête de i vers j , et où chaque arête est étiquetée par un réel p_{ij} compris entre 0 et 1, la somme des poids des arêtes issues d'un même sommet étant égale à 1.

• **Point méthode pour construire un graphe probabiliste**

On comptabilise les différents états.

1. S'il y en a N , on représente une figure géométrique à N sommets.
2. On dessine des arcs orientés entre les sommets avec un poids qui correspond à la probabilité de passage d'un état à un autre (éventuellement à lui-même).
3. On vérifie que la somme des poids des arcs issus d'un même sommet est égale à 1.

De même qu'à un graphe (orienté ou non), on associe une matrice d'adjacence A , dont le terme a_{ij} compte le nombre d'arêtes joignant le sommet i au sommet j , on peut associer à un graphe probabiliste une matrice qui décrit les probabilités de transition :

Définition 11.17 (Matrice de transition).

Étant donné un graphe probabiliste à N sommets, on appelle *matrice de transition* associée la matrice carrée $M = (p_{ij})$ à N lignes et N colonnes, dont le coefficient p_{ij} est l'étiquette de l'arête orientée de i vers j si elle existe (c'est-à-dire la probabilité de transition de i à j), et 0 sinon.

EXERCICE 11.1.

On considère un pentagone régulier dont les sommets sont numérotés de 1 à 5. Un promeneur saute à chaque instant d'un sommet à l'un de ses voisins de manière équiprobable.

Représenter le graphe probabiliste associé ainsi que sa matrice de transition.

11.2.3 Chaînes de Markov

Si on lance un grand nombre de fois une pièce de monnaie, on obtient la plupart du temps autant de piles que de faces. Cela se traduit en mathématiques par le théorème appelé loi des grands nombres. Cette loi exprime que, lorsque les lancers sont indépendants les uns des autres, la fréquence empirique de réalisation d'un événement va tendre vers la probabilité théorique de réalisation de cet événement.

Sous cette forme, la loi des grands nombres est loin de couvrir tous les phénomènes que l'on peut modéliser par une expérience aléatoire. Prenons l'exemple du temps qu'il fait. Bien sûr, la météo du jour dépend de celle de la veille. Mais on peut dire qu'elle ne dépend pas de celle d'il y a un an ! La question de savoir si on peut étendre la loi des grands nombres à une suite d'expériences aléatoires qui dépendent les unes des autres, mais pas trop, a été résolue par le mathématicien Markov en 1902, en introduisant ce qu'on appelle maintenant les chaînes de Markov.

Définition 11.18 (Espace des états).

On considère dans la suite un ensemble E , appelé espace des états et une suite (X_n) de variables aléatoires à valeurs dans E .

Exemple 11.5. On considère trois points A, B, C . Un pion est placé en A au début de la partie (qui correspond à l'instant $t = 0$), et il se déplace suivant les règles suivantes :

- si le pion est en A à l'instant t , à l'instant $t + 1$ il se trouve en A, B ou C avec une probabilité de $\frac{1}{3}$;
- si le pion est en B à l'instant t , à l'instant $t + 1$ il se trouve en B ou C avec une probabilité de $\frac{1}{2}$;
- dès que le pion est en C , il n'en bouge plus. On s'intéresse à l'évolution de la position du pion au cours du temps.

Alors $E = \{A, B, C\}$, X_n est la position du pion à l'instant n .

Définition 11.19 (🔗 Chaînes de Markov).

On appelle chaîne de Markov, toute suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires (définies sur le même espace probabilisé) à valeurs dans un espace d'états E telles que pour tout entier naturel n et pour tout $(e_0, e_1, \dots, e_{n+1}) \in E^{n+2}$, tels que $P([X_0 = e_0] \cap [X_1 = e_1] \cdots \cap [X_n = e_n]) > 0$,

$$P_{[X_0=e_0] \cap [X_1=e_1] \cap [X_2=e_2] \cdots \cap [X_n=e_n]}(X_{n+1} = e_{n+1}) = P_{[X_n=e_n]}(X_{n+1} = e_{n+1}).$$

La loi de X_{n+1} sachant $X_0, X_1, \dots, X_n$ est égale à la loi de X_{n+1} sachant X_n .

Remarque. Les chaînes de Markov sont donc un exemple simple de suites de variables aléatoires non indépendantes.

Si n désigne le temps présent, la propriété de Markov signifie que, pour connaître le futur (l'état de X_{n+1}) conditionnellement à ce qui s'est passé avant l'instant n ($X_0, \dots, X_n$), il suffit de connaître le présent (X_n) ce qui est souvent résumé en une phrase ressemblant à « le futur est indépendant du passé conditionnellement au présent ». D'une certaine façon, le processus est sans mémoire.

C'est Markov lui-même qui a le premier appliqué ses chaînes à un domaine autre que les mathématiques. Il s'en est servi pour analyser les successions de lettres dans l'alphabet russe. Le même type d'applications se retrouve aujourd'hui dans les téléphones portables : les chaînes de Markov sont mises à contribution par les logiciels d'écriture de SMS, qui essaient de deviner ce que vous allez écrire. Plus sérieusement, les chaînes de Markov sont devenues un outil très employé, pour l'analyse de l'ADN, la prédiction de l'évolution des épidémies, ou pour classer l'importance des pages internet....

Définitions 11.20 (🔗 Chaîne de Markov homogène).

On dit que la chaîne est dans l'état i au temps n si l'événement $[X_n = i]$ est réalisé. L'ensemble E est l'ensemble des états de la chaîne.

$\forall (i, j) \in E^2$, les probabilités $P_{[X_n=i]}(X_{n+1} = j)$ notées $p_{i,j}(n)$ sont appelées probabilités de transition.

Une chaîne de Markov est dite homogène lorsque les probabilités $p_{i,j}(n)$ ne dépendent pas de n mais seulement de i et j , elles sont alors notées $p_{i,j}$ et on a :

$$\forall (i, j) \in E^2, P_{[X_n=i]}(X_{n+1} = j) = P_{[X_0=i]}(X_1 = j) = p_{i,j}.$$

Remarque. La propriété d'homogénéité signifie que la probabilité de passer d'un état à un autre ne dépend pas de l'instant auquel on est. On peut montrer que, quitte à changer d'espace des états, on peut se contenter d'étudier des chaînes homogènes.

On peut demander de justifier au concours si toute suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires définit une chaîne de Markov homogène.

La phrase type est alors que d'une part :

D'après l'énoncé, à un instant donné n , ne dépend que de sa situation au temps n . (X_n) est donc une chaîne de Markov.

D'autre part :

La règle pour passer de n à $n+1$, est la même indépendamment de l'instant : la suite (X_n) définit bien une chaîne de Markov homogène.

Voir exos TP5 pour des exemples.

Exemple 11.6 (Evolution d'un titre boursier au cours du temps). Dans une bourse de valeurs, un titre donné peut augmenter, rester stable, ou baisser. Dans un modèle mathématique, on considère que :

- Le titre est stable à l'instant 0.
- Si à l'instant n , le titre augmente, alors à l'instant $n+1$, soit il augmente, avec une probabilité de $2/3$, soit il baisse ou reste stable, et ceci avec la même probabilité.
- Si à l'instant n , le titre reste stable, alors à l'instant $n+1$, soit il reste stable, avec une probabilité

de $2/3$, soit il augmente ou baisse, et ceci avec la même probabilité.

• Si à l'instant n , le titre baisse, alors à l'instant $n + 1$, soit il baisse, avec une probabilité de $2/3$, soit il augmente ou reste stable, et ceci avec la même probabilité.

On note 1 l'état correspondant à l'état "le titre augmente", 2 celui correspondant à l'état "le titre reste stable" et 3 celui correspondant à "le titre baisse". Pour tout entier n , X_n représente le titre boursier à l'instant n .

Justifions que la suite (X_n) définit une chaîne de Markov homogène.

D'après l'énoncé, à un instant donné n , l'évolution du titre ne dépend que de sa situation au temps n .

(X_n) est donc une chaîne de Markov. La règle de l'évolution pour passer de n à $n + 1$ (suivant que le titre augmente, reste stable ou baisse), est la même indépendamment de l'instant : la suite (X_n) définit bien une chaîne de Markov homogène.

On veut étudier l'évolution d'un tel système (chaîne homogène à valeurs dans E fini) au cours du temps; on note V_n le vecteur ligne à N éléments dont l'élément d'ordre j est la probabilité que le système se trouve à l'instant n dans l'état j (Attention aux indices! on a ici une suite de vecteurs lignes, avec chacun N composantes!).

Définition 11.21 (✎ n^e état probabiliste).

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène à N états notés $1, 2, \dots, N$.

Pour tout entier naturel n , le n^e état probabiliste de la chaîne de Markov est la matrice ligne définie par :

$V_n = (P(X_n = 1), \dots, P(X_n = N))$. V_0 est appelé état initial.

Définition 11.22 (✎ Matrice de transition).

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène à N états notés $1, 2, \dots, N$ de probabilités transitionnelles $p_{i,j} = P_{[X_n=i]}(X_{n+1} = j)$, on appelle matrice de transition de la chaîne de Markov, la matrice carrée d'ordre N suivante :

$$M = \begin{pmatrix} p_{1,1} & p_{1,2} & \dots & p_{1,N} \\ p_{2,1} & p_{2,2} & \dots & p_{2,N} \\ \dots & \dots & \dots & \dots \\ p_{N,1} & p_{N,2} & \dots & p_{N,N} \end{pmatrix}$$

Remarque. Les coefficients de cette matrice sont positifs et la somme des coefficients de chaque ligne est égale à 1. Une telle matrice est dite stochastique.

Définition 11.23 (✎ Graphe associé).

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène de matrice de transition $M = (p_{i,j}) \in \mathbb{M}_N(\mathbb{R})$. On appelle graphe associé à M , le graphe à N sommets tels que le poids de l'arc orienté qui permet de passer du sommet i au sommet j est $p_{i,j}$.

• **Point méthode pour construire un graphe associé à une matrice de transition**

Soit $M = (p_{i,j})$ à N lignes et N colonnes,

1. Pour dessiner le graphe associé à M , on place les N sommets.
2. Puis on place le poids $p_{i,j}$ sur l'arc orienté qui permet de passer du sommet i au sommet j .

EXERCICE 11.2.

On donne la matrice $M = \begin{pmatrix} 0,4 & 0,6 \\ 0,7 & 0,3 \end{pmatrix}$. Dessiner le graphe associé à M .

Théorème 11.8 (🔗).

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène à N états notés $1, 2, \dots, N$ et de matrice de transition $M = (p_{i,j}) \in \mathbb{M}_N(\mathbb{R})$, alors pour tout entier naturel n , on a :

$\forall j \in [1, N], P(X_{n+1} = j) = \sum_{i=1}^N P(X_n = i) p_{i,j}$, ce qui s'écrit $V_{n+1} = V_n \times M$ (où V_n désigne le n^e état probabiliste de la chaîne).

Remarque. Le premier résultat s'obtient grâce à la formule des probabilités totales.

Le second n'est que la traduction matricielle du premier.

On en déduit immédiatement :

Théorème 11.9 (🔗).

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène à N états notés $1, 2, \dots, N$ et de matrice de transition $M = (p_{i,j}) \in \mathbb{M}_N(\mathbb{R})$, le n^e état probabiliste de la chaîne est donné par :

Pour tout entier $n > 0$, on a : $V_n = V_0 \times M^n$.

Preuve. A savoir faire par récurrence. ◇

Cette formule, qui permet de calculer la répartition de probabilités au temps n si on la connaît au temps 0, est fondamentale pour tous les exercices.

EXERCICE 11.3.

On considère un mobile qui se déplace par sauts sur les points O, A, B selon la règle suivante :

- S'il est en O à l'instant n , il sera de façon équiprobable, soit en A , soit en B à l'instant $n + 1$.
- S'il est en A à l'instant n , il retourne de façon équiprobable, soit en O , soit il reste en A à l'instant $n + 1$.
- S'il est en B à l'instant n , il passe en O à l'instant $n + 1$ avec la probabilité 1 (le point B est dit « réfléchissant »).

On dit que le mobile est dans l'état 1 s'il se trouve en A , dans l'état 2 s'il se trouve en O et dans l'état 3 s'il se trouve en B . On notera, pour tout entier naturel n , X_n l'état dans lequel se trouve le mobile à l'instant n .

On admet que $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène.

1. Représenter le graphe probabiliste (=graphe associé à cette chaîne).
2. Déterminer la matrice de transition M .

Un cas particulièrement intéressant est celui où la répartition de probabilité est stable au cours du temps.

11.2.4 Etats stables d'une chaîne de Markov homogène**Définition 11.24** (🔗 Etat stable).

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène à N états notés $1, 2, \dots, N$ et de matrice de transition M .

On appelle état stable de la chaîne de Markov ou loi stationnaire, une loi de probabilité définie par vecteur ligne $\pi = (x_1 \dots x_N)$ à N composantes tel que $\pi = \pi \times M$ avec $\sum_{i=1}^N x_i = 1$ et $\forall i \in \{1, \dots, N\}, x_i \geq 0$.

Remarque. Si l'état initial X_0 a pour loi de probabilité π , alors X_n a aussi pour loi de probabilité π , indépendamment de n .

Propriété 11.10 (État stable et vecteur propre).

Un état stable π s'il existe, d'une chaîne de Markov de matrice de transition M est tel que ${}^t\pi$ est un vecteur propre de tM associé à la valeur propre 1.

Preuve. Un état stable vérifie $\pi = \pi \times M$, alors en transposant, on obtient : ${}^t\pi = {}^tM {}^t\pi$, ce qui prouve bien que ${}^t\pi$ est un vecteur propre de tM associé à la valeur propre 1. De plus, on le choisit de telle sorte que ses composantes soient positives et que leur somme est égale à 1. $\diamond$

Théorème 11.11 (Loi limite).

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène. La loi limite de $(X_n)_{n \in \mathbb{N}}$, si elle existe est un état stable de la chaîne.

Remarque. Le théorème 11.11 nous explique que si l'on calcule la limite quand n tend vers $+\infty$ de $P(X_n = j)$, pour tout $j \in \llbracket 1, N \rrbracket$, on définit l'état stable.

Remarquons que la recherche d'un état stable n'est pas difficile en pratique : il s'agit de résoudre l'équation $\pi = \pi \times M$, et d'en chercher une solution satisfaisant $\sum_{i=1}^N x_i = 1$, ce qui se fait sans problème pour une matrice M donnée.

On peut interpréter cette équation de la manière suivante : pour obtenir un état stable, il faut que toutes les transitions s'équilibrent ; si l'on considère le problème comme une évolution de population, il faut que, pour tout état i , la quantité de personnes qui quittent l'état i à chaque étape soit égale à la quantité de personnes qui y arrivent.

Ce qui est moins évident, c'est qu'un tel état existe, et qu'il soit unique. Ce qui se passe quand on ne part pas d'un état stable n'est pas évident non plus. Nous allons l'étudier en détail dans le cas d'un système à deux états.

EXERCICE 11.4.

On continue l'étude de la chaîne de Markov de l'exercice 11.3.1.

Déterminer les états stables possibles.

EXERCICE 11.5.

Un individu vit dans un milieu où il est susceptible d'attraper une maladie par piqure d'insecte. Il peut être dans l'un des trois états suivants : immunisé (I), malade (M), sain, c'est-à-dire non malade et non immunisé, (S).

D'un mois à l'autre, son état peut changer selon les règles suivantes :

- étant immunisé, il peut le rester avec une probabilité 0,9 ou passer à l'état S avec une probabilité 0,1 ;
- étant dans l'état S , il peut le rester avec une probabilité 0,5 ou passer à l'état M avec une probabilité 0,5 ;
- étant malade, il peut le rester avec une probabilité 0,2 ou passer à l'état I avec une probabilité 0,8.

1. Tracer un graphe probabiliste pour décrire cette situation et écrire la matrice de transition.
2. Calculer l'état de probabilité de l'individu au bout de trois mois, d'un an, de 2 ans pour chacune des situations suivantes :
 - au départ, il est immunisé ;
 - au départ, il est non malade et non immunisé ;

- au départ, il est malade.

3. On note $(i \ m \ s)$ l'état stable du système. Déterminer les états stables possibles.

EXERCICE 11.6.

Un mobile se déplace aléatoirement dans l'ensemble des sommets d'un triangle ABC de la façon suivante :

- Le mobile se trouve sur le point A à l'instant 0.
- Si à l'instant n , il est sur l'un quelconque des sommets, alors à l'instant $(n + 1)$, soit il y reste, avec une probabilité de $2/3$, soit il se place sur l'un des deux autres sommets et ceci avec la même probabilité.

1. Dessiner le graphe probabiliste associé à cette chaîne (on notera 1 l'état correspondant au sommet A , 2 celui correspondant au sommet B et 3 celui correspondant au sommet C).

2. Donner la matrice de transition M .

3. On donne $P = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ et $Q = \begin{pmatrix} -1 & 2 & -1 \\ -1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$.

(a) Calculer PQ , en déduire que P est inversible et donner P^{-1} .

(b) Montrer que $P^{-1}MP$ est une matrice diagonale D à déterminer.

(c) En déduire pour tout entier naturel n , l'expression de M^n .

4. Pour tout entier naturel n , déterminer le n^e état probabiliste V_n .

5. Montrer que les composantes de V_n ont pour limites respectives $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$. Vérifier que le vecteur $(\frac{1}{3} \ \frac{1}{3} \ \frac{1}{3})$ est un état stable.

EXERCICE 11.7.

Chaque année, une association de cyclotourisme prépare de nouveaux circuits. Pour satisfaire ses nombreux membres, elle élabore des circuits de différents niveaux : « niveau facile », « niveau moyen » et « niveau difficile ».

Au premier janvier 2010, l'association a fait son bilan :

- 20 % de ses adhérents ont choisi le niveau facile, noté A
- 70 % de ses adhérents ont choisi le niveau moyen, noté B
- 10 % de ses adhérents ont choisi le niveau difficile, noté C

Pour répondre aux attentes des adhérents et les fidéliser sur le long terme, une enquête est effectuée. Il s'avère que, d'une année à l'autre :

- parmi les adhérents ayant choisi le niveau A, 40 % restent à ce niveau et 60 % passent au niveau B,
- parmi les adhérents ayant choisi le niveau B, 70 % restent à ce niveau et 20 % reviennent au niveau A et les autres passent au niveau C,
- parmi les adhérents ayant choisi le niveau C, 85 % restent à ce niveau et les autres reviennent au niveau B.

On note :

- A l'état « l'adhérent a choisi le niveau A »,
- B l'état « l'adhérent a choisi le niveau B »,
- C l'état « l'adhérent a choisi le niveau C ».

Pour n entier naturel positif ou nul, on note $P_n = (a_n \ b_n \ c_n)$ la matrice ligne donnant l'état probabiliste de la répartition dans les différents niveaux (indiqués dans l'ordre donné dans l'énoncé), au premier janvier de l'année 2010 + n .

On se décide se baser uniquement sur ces résultats pour prévoir l'évolution de la répartition à partir du premier janvier 2010 (on néglige donc les nouveaux abonnés et les départs).

1. Donner P_0 .
2. Représenter cette situation par un graphe probabiliste de sommets A, B et C.
3. Donner la matrice M de transition associée à ce graphe, en respectant l'ordre alphabétique des sommets.
4. Une seule des trois matrices Q , R , T ci-dessous correspond à l'état probabiliste stable.

$$Q = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \quad R = \begin{pmatrix} \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \end{pmatrix} \quad T = \begin{pmatrix} \frac{1}{5} & \frac{4}{5} & 0 \end{pmatrix}$$

Le président de l'association affirme que 50 % des adhérents choisiront après un certain nombre d'années le niveau B. Cette affirmation est-elle correcte ?

• **Point méthode pour étudier une chaîne de Markov**

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène à N états notés $1, 2, \dots, N$.

1. On peut demander de représenter le graphe probabiliste associé à cette chaîne (voir point méthode 1). On définit les états de la chaîne par des entiers successifs en démarrant à 1.
2. On peut demander de définir la matrice de transition M : il suffit de reporter sur la ligne i les poids des arcs issus du sommet qui représente l'état i en suivant l'ordre naturel des états.
3. Pour déterminer le n^e état probabiliste de la chaîne, ie la loi de probabilité de X_n , c'est un calcul algébrique et il suffit de se laisser guider.
4. Pour déterminer un état stable s'il existe, cela revient à chercher la loi limite de $(X_n)_{n \in \mathbb{N}}$ si elle existe ou trouver parmi les vecteurs propres de ${}^t M$ des vecteurs dont les composantes sont positives et de somme égale à 1.
Ou on cherche π un vecteur ligne (composé de réels positifs dont la somme vaut 1) et tel que $\pi M = \pi$.

11.3 Un cas particulier : les graphes probabilistes à 2 états

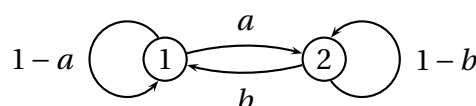
On suppose qu'il n'y a que deux états, notés 1 et 2. On note U_n (resp. V_n) la probabilité qu'à l'instant n , le système se trouve dans l'état 1 (resp. 2).

Pour tout entier n , on note X_n le vecteur-ligne à deux colonnes, $X_n = (U_n \ V_n)$. Remarquons que l'on a toujours $U_n + V_n = 1$.

On note a la probabilité de transition de l'état 1 à l'état 2, c'est-à-dire la probabilité que le système passe à l'état 2 à l'étape $n+1$ sachant que le système est à l'état 1 à l'étape n et b la probabilité de transition de l'état 2 à l'état 1.

Les probabilités de demeurer dans l'état 1 ou dans l'état 2 sont donc $1-a$ et $1-b$.

Le système peut donc être représenté par le graphe probabiliste suivant :



et la matrice correspondante est :

$$M = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

Les nombres a et b sont tous deux compris entre 0 et 1.

Il y a quelques **cas triviaux** que l'on peut traiter à part :

- Si a et b sont nuls, M est la matrice identité, et **tous les états sont stables**. Ce n'est pas étonnant, puisqu'il est impossible de changer!
- Si l'un des deux seulement est nul, par exemple a , on voit que l'on finit toujours par arriver dans l'état 1. C'est le cas d'une population qui ne se renouvelle pas, et dont les individus peuvent être dans deux états, vivants (état 2) ou morts (état 1) : à long terme, la population ne sera plus composée que de morts. On parle alors d'état absorbant (pour l'état 1). Il y a donc **un état stable**.
- Enfin, si les deux coefficients a et b sont égaux à 1, le système clignote : il oscille sans se stabiliser entre l'état 1 et l'état 2, et se retrouve tous les deux coups dans le même état. **Il n'y a pas d'état stable**.

Ces cas particuliers étant faciles à étudier, et peu intéressants, **on supposera désormais que a et b sont strictement positifs, et ne sont pas tous deux égaux à 1.**

On peut démontrer alors le résultat général suivant :

Théorème 11.12.

Considérons un graphe probabiliste à deux états, de matrice de transition $M = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$ telle que $0 < a < 1$ et $0 < b < 1$. Alors, le système admet un unique état stable, indépendamment de l'état initial.

Remarque. On peut généraliser, sous certaines conditions à un graphe à n états :

Théorème 11.13.

Soit un graphe probabiliste à n états, de matrice de transition M . S'il existe une puissance M^k de M dont tous les coefficients sont strictement positifs, alors il existe un seul état stable X , vérifiant $X = XM$, et quel que soit l'état initial, le système converge exponentiellement vite vers l'état stable.

On l'admettra.

11.3.1 Exercices

EXERCICE 11.8.

Un théâtre propose deux types d'abonnements pour une année : un abonnement A donnant droit à six spectacles ou un abonnement B donnant droit à trois spectacles.

On considère un groupe de 2 500 personnes qui s'abonnent tous les ans. n étant un entier naturel, on note :

a_n la probabilité qu'une personne ait choisi un abonnement A l'année n ;

b_n la probabilité qu'une personne ait choisi un abonnement B l'année n ;

P_n la matrice $(a_n \ b_n)$ traduisant l'état probabiliste à l'année n .

Tous les ans 85 % des personnes qui ont choisi l'abonnement A et 55 % des personnes qui ont choisi l'abonnement B conservent ce type d'abonnement l'année suivante. Les autres personnes changent d'abonnement.

1. On suppose que, l'année zéro, 1 500 personnes ont choisi l'abonnement A et 1 000 l'abonnement B.

En déduire l'état initial $P_0 = (a_0 \ b_0)$.

2. (a) Tracer un graphe probabiliste traduisant les données de l'énoncé.
 (b) Déterminer la matrice de transition M de ce graphe.
 (c) Déterminer le nombre d'abonnés pour chaque type d'abonnement l'année n .
3. Prouver qu'il existe un unique état stable, le déterminer.
 Interpréter le résultat précédent en terme de nombre d'abonnements de type A et de type B .

EXERCICE 11.9.

Étude de l'évolution météorologique d'un jour à l'autre dans une localité.

- S'il fait sec aujourd'hui, alors il fera encore sec demain avec la probabilité $\frac{5}{6}$, donc il fera humide demain avec la probabilité $\frac{1}{6}$.
- S'il fait humide aujourd'hui, alors il fera encore humide demain avec la probabilité $\frac{2}{3}$.

Nous sommes dimanche et il fait sec. On s'intéresse à l'évolution météorologique des jours suivants.

1. Soit n un entier naturel, on note :
 s_n la probabilité pour que le jour n , il fasse sec ;
 h_n la probabilité pour que le jour n , il fasse humide ;
 T_n la matrice $(s_n \ h_n)$ traduisant l'état probabiliste du temps le jour n . Déterminer une relation entre s_n et h_n .
2. (a) Si le premier dimanche est le jour correspondant à $n = 0$, donner la matrice associée à l'état initial du temps.
 (b) Décrire l'évolution de cet état à l'aide d'un graphe probabiliste.
3. Donner la matrice M de transition associée à ce graphe.
 (a) Déterminer M^2 .
 (b) Quelle sera la situation météorologique le mardi suivant?
4. (a) Déterminer l'état stable associé à l'évolution météorologique.
 (b) En déduire, à long terme, la probabilité qu'il pleuve un certain jour.

EXERCICE 11.10 (L'allumeur de réverbères).

Chaque matin, l'allumeur de réverbère du Petit Prince change l'état de sa planète avec une probabilité 0,75. Au jour 0, le réverbère est éteint.

1. Décrire cette situation à l'aide d'un graphe probabiliste.
2. Soit M la matrice de transition associée à ce graphe.

Vérifier que : $M = N - \frac{1}{2}R$, où $N = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ et $R = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$

3. Calculer N^2 , R^2 , NR et RN puis en déduire M^n , pour n entier naturel.
4. Au jour 0, le réverbère est allumé (respectivement éteint). Calculer la probabilité p_n (respectivement p'_n) que le réverbère soit allumé. (respectivement éteint) au n -ième matin.
5. On constate que sur une durée de 6 jours, pour une population de 500 réverbères, une stabilisation assez rapide de la situation suivante : un réverbère sur 2 est allumé. Comment peut-on expliquer cette observation?

Feuille d'exercices n°11

CHAÎNE DE MARKOV

Exercice 1. On considère un pentagone régulier dont les sommets sont numérotés de 1 à 5. Un promeneur saute à chaque instant d'un sommet à l'un de ses voisins de manière équiprobable. Représenter le graphe probabiliste associé ainsi que sa matrice de transition.

Exercice 2. On donne la matrice $M = \begin{pmatrix} 0,4 & 0,6 \\ 0,7 & 0,3 \end{pmatrix}$. Dessiner le graphe associé à M .

Exercice 3. On considère un mobile qui se déplace par sauts sur les points O, A, B selon la règle suivante :

- S'il est en O à l'instant n , il sera de façon équiprobable, soit en A , soit en B à l'instant $n + 1$.
- S'il est en A à l'instant n , il retourne de façon équiprobable, soit en O , soit il reste en A à l'instant $n + 1$.
- S'il est en B à l'instant n , il passe en O à l'instant $n + 1$ avec la probabilité 1 (le point B est dit « réfléchissant »).

On dit que le mobile est dans l'état 1 s'il se trouve en A , dans l'état 2 s'il se trouve en O et dans l'état 3 s'il se trouve en B . On notera, pour tout entier naturel n , X_n l'état dans lequel se trouve le mobile à l'instant n .

On admet que $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov homogène.

1. Représenter le graphe probabiliste (=graphe associé à cette chaîne).
2. Déterminer la matrice de transition M .
3. Déterminer les états stables possibles.

Exercice 4. Un individu vit dans un milieu où il est susceptible d'attraper une maladie par piqûre d'insecte. Il peut être dans l'un des trois états suivants : immunisé (I), malade (M), sain, c'est-à-dire non malade et non immunisé, (S).

D'un mois à l'autre, son état peut changer selon les règles suivantes :

- étant immunisé, il peut le rester avec une probabilité 0,9 ou passer à l'état S avec une probabilité 0,1;
- étant dans l'état S , il peut le rester avec une probabilité 0,5 ou passer à l'état M avec une probabilité 0,5;
- étant malade, il peut le rester avec une probabilité 0,2 ou passer à l'état I avec une probabilité 0,8.

1. Tracer un graphe probabiliste pour décrire cette situation et écrire la matrice de transition.

2. Calculer l'état de probabilité de l'individu au bout de trois mois, d'un an, de 2 ans pour chacune des situations suivantes :
 - au départ, il est immunisé;
 - au départ, il est non malade et non immunisé;
 - au départ, il est malade.
3. On note $(i \ m \ s)$ l'état stable du système. Déterminer les états stables possibles.

Exercice 5. Un mobile se déplace aléatoirement dans l'ensemble des sommets d'un triangle ABC de la façon suivante :

- Le mobile se trouve sur le point A à l'instant 0.
- Si à l'instant n , il est sur l'un quelconque des sommets, alors à l'instant $(n + 1)$, soit il y reste, avec une probabilité de $2/3$, soit il se place sur l'un des deux autres sommets et ceci avec la même probabilité.

1. Dessiner le graphe probabiliste associé à cette chaîne (on notera 1 l'état correspondant au sommet A , 2 celui correspondant au sommet B et 3 celui correspondant au sommet C).
2. Donner la matrice de transition M .
3. On donne $P = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ et $Q = \begin{pmatrix} -1 & 2 & -1 \\ -1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$.
 - (a) Calculer PQ , en déduire que P est inversible et donner P^{-1} .
 - (b) Montrer que $P^{-1}MP$ est une matrice diagonale D à déterminer.
 - (c) En déduire pour tout entier naturel n , l'expression de M^n .
4. Pour tout entier naturel n , déterminer le n^e état probabiliste V_n .
5. Montrer que les composantes de V_n ont pour limites respectives $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$. Vérifier que le vecteur $(\frac{1}{3} \ \frac{1}{3} \ \frac{1}{3})$ est un état stable.

Exercice 6. Un théâtre propose deux types d'abonnements pour une année : un abonnement A donnant droit à six spectacles ou un abonnement B donnant droit à trois spectacles.

On considère un groupe de 2 500 personnes qui s'abonnent tous les ans. n étant un entier naturel, on note :

a_n la probabilité qu'une personne ait choisi un abonnement A l'année n ;

b_n la probabilité qu'une personne ait choisi un abonnement B l'année n ;

P_n la matrice $(a_n \ b_n)$ traduisant l'état probabiliste à l'année n .

Tous les ans 85 % des personnes qui ont choisi l'abonnement A et 55 % des personnes qui ont choisi l'abonnement B conservent ce type d'abonnement l'année suivante. Les autres personnes changent d'abonnement.

1. On suppose que, l'année zéro, 1 500 personnes ont choisi l'abonnement A et 1 000 l'abonnement B.
En déduire l'état initial $P_0 = (a_0 \ b_0)$.
2. (a) Tracer un graphe probabiliste traduisant les données de l'énoncé.
(b) Déterminer la matrice de transition M de ce graphe.
(c) Déterminer le nombre d'abonnés pour chaque type d'abonnement l'année n .
3. Prouver qu'il existe un unique état stable, le déterminer.
Interpréter le résultat précédent en terme de nombre d'abonnements de type A et de type B.

Exercice 7. *Étude de l'évolution météorologique d'un jour à l'autre dans une localité.*

- S'il fait sec aujourd'hui, alors il fera encore sec demain avec la probabilité $\frac{5}{6}$, donc il fera humide demain avec la probabilité $\frac{1}{6}$.
- S'il fait humide aujourd'hui, alors il fera encore humide demain avec la probabilité $\frac{2}{3}$.

Nous sommes dimanche et il fait sec. On s'intéresse à l'évolution météorologique des jours suivants.

1. Soit n un entier naturel, on note :
 s_n la probabilité pour que le jour n , il fasse sec;
 h_n la probabilité pour que le jour n , il fasse humide;
 T_n la matrice $(s_n \ h_n)$ traduisant l'état probabiliste du temps le jour n . Déterminer une relation entre s_n et h_n .
2. (a) Si le premier dimanche est le jour correspondant à $n = 0$, donner la matrice associée à l'état initial du temps.
 (b) Décrire l'évolution de cet état à l'aide d'un graphe probabiliste.
3. Donner la matrice M de transition associée à ce graphe.
 (a) Déterminer M^2 .
 (b) Quelle sera la situation météorologique le mardi suivant?
4. (a) Déterminer l'état stable associé à l'évolution météorologique.
 (b) En déduire, à long terme, la probabilité qu'il pleuve un certain jour.

Exercice 8. Chaque année, une association de cyclotourisme prépare de nouveaux circuits. Pour satisfaire ses nombreux membres, elle élabore des circuits de différents niveaux : « niveau facile », « niveau moyen » et « niveau difficile ».

Au premier janvier 2010, l'association a fait son bilan :

- 20 % de ses adhérents ont choisi le niveau facile, noté A
- 70 % de ses adhérents ont choisi le niveau moyen, noté B
- 10 % de ses adhérents ont choisi le niveau difficile, noté C

Pour répondre aux attentes des adhérents et les fidéliser sur le long terme, une enquête est effectuée. Il s'avère que, d'une année à l'autre :

- parmi les adhérents ayant choisi le niveau A, 40 % restent à ce niveau et 60 % passent au niveau B,
- parmi les adhérents ayant choisi le niveau B, 70 % restent à ce niveau et 20 % reviennent au niveau A et les autres passent au niveau C,
- parmi les adhérents ayant choisi le niveau C, 85 % restent à ce niveau et les autres reviennent au niveau B.

On note :

- A l'état « l'adhérent a choisi le niveau A »,
- B l'état « l'adhérent a choisi le niveau B »,
- C l'état « l'adhérent a choisi le niveau C ».

Pour n entier naturel positif ou nul, on note $P_n = (a_n \ b_n \ c_n)$ la matrice ligne donnant l'état probabiliste de la répartition dans les différents niveaux (indiqués dans l'ordre donné dans l'énoncé), au premier janvier de l'année 2010 + n .

On se décide se baser uniquement sur ces résultats pour prévoir l'évolution de la répartition à partir du premier janvier 2010 (on néglige donc les nouveaux abonnés et les départs).

1. Donner P_0 .
2. Représenter cette situation par un graphe probabiliste de sommets A, B et C.
3. Donner la matrice M de transition associée à ce graphe, en respectant l'ordre alphabétique des sommets.
4. Une seule des trois matrices Q , R , T ci-dessous correspond à l'état probabiliste stable.

$$Q = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \quad R = \begin{pmatrix} \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \end{pmatrix} \quad T = \begin{pmatrix} \frac{1}{5} & \frac{4}{5} & 0 \end{pmatrix}$$

Le président de l'association affirme que 50 % des adhérents choisiront après un certain nombre d'années le niveau B. Cette affirmation est-elle correcte ?

Exercice 9. [L'allumeur de réverbères]

Chaque matin, l'allumeur de réverbère du Petit Prince change l'état de sa planète avec une probabilité 0,75. Au jour 0, le réverbère est éteint.

1. Décrire cette situation à l'aide d'un graphe probabiliste.
2. Soit M la matrice de transition associée à ce graphe.
Vérifier que : $M = N - \frac{1}{2}R$, où $N = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ et $R = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$
3. Calculer N^2 , R^2 , NR et RN puis en déduire M^n , pour n entier naturel.
4. Au jour 0, le réverbère est allumé (respectivement éteint). Calculer la probabilité p_n (respectivement p'_n) que le réverbère soit allumé. (respectivement éteint) au n -ième matin.
5. On constate que sur une durée de 6 jours, pour une population de 500 réverbères, une stabilisation assez rapide de la situation suivante : un réverbère sur 2 est allumé. Comment peut-on expliquer cette observation ?

Exercice 10. Une urne A contient deux jetons numérotés 0 et une urne B contient deux jetons numérotés 1. On tire un jeton dans chaque urne et on les échange (le jeton provenant de A est placé dans B , celui prélevé dans B est remis dans A). On procède ainsi à n échanges ($n \in \mathbb{N}$). On note X_n la somme aléatoire des numéros des jetons alors contenus dans A . Ainsi X_0 est la variable aléatoire constamment nulle. Pour n entier naturel, on pose : $p_n = P(X_n = 0)$, $q_n = P(X_n = 1)$ et $r_n = P(X_n = 2)$.

1. Pour n entier naturel, exprimer p_{n+1} , q_{n+1} et r_{n+1} en fonction de p_n , q_n et r_n .
2. Pour n entier naturel, exprimer q_{n+2} en fonction de q_{n+1} et q_n . En déduire qu'il existe deux réels λ et μ (que l'on déterminera) tels que : $\forall n \in \mathbb{N}, q_n = \lambda + \mu \left(-\frac{1}{2}\right)^n$.
3. Pour n entier naturel, déterminer la loi de X_n et calculer $E(X_n)$.
4. Déterminer $\lim_{n \rightarrow +\infty} p_n$, $\lim_{n \rightarrow +\infty} q_n$, $\lim_{n \rightarrow +\infty} r_n$, et $\lim_{n \rightarrow +\infty} E(X_n)$.

Exercice 11. Partie A

On considère les matrices suivantes :

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 4 & 1 \\ 4 & 0 & 4 \end{pmatrix}, P = \begin{pmatrix} 1 & 1 & -1 \\ 4 & 1 & 0 \\ 4 & -2 & 1 \end{pmatrix}, B = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, J = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, K = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

1. Montrer que P est inversible et calculer P^{-1} .
2. Calculer $P^{-1}AP$.
3. Calculer JK, KJ, J^n, K^n pour $n \in \mathbb{N}$. Écrire B en fonction de J et K , en déduire B^n .

4. En déduire B^n .

Partie B

Un jeton est placé au hasard, c'est à dire avec équiprobabilité, sur l'une des 3 cases d'une planchette numérotées de gauche à droite : 1, 2 et 3. Le déplacement du jeton sur la planchette est régi par l'expérience suivante :

- On tire une boule d'une urne contenant 4 boules blanches et une boule noire :
 - si la boule tirée est blanche, on déplace le jeton de 2 cases vers la droite, ou, si ce déplacement est impossible, on laisse le jeton où il se trouve.
 - si la boule tirée est noire, on déplace le jeton de 1 case vers la gauche, ou, si ce déplacement est impossible, on laisse le jeton où il se trouve.
- On remet la boule tirée dans l'urne.

Pour $n \in \mathbb{N}^*$, on note X_n la variable aléatoire indiquant le numéro de la case où se trouve le jeton après le n^{eme} tirage dans l'urne. X_0 désigne la variable aléatoire indiquant le numéro de la case où se trouve le jeton avant le premier tirage.

1. Préciser la loi de X_0 . Calculer $E(X_0)$ et $V(X_0)$.
2. Décrire la situation par un graphe probabiliste.
3. Pour $n \in \mathbb{N}$, on note : $a_n = P(X_n = 1)$, $b_n = P(X_n = 2)$, $c_n = P(X_n = 3)$ et $U_n = (a_n \ b_n \ c_n)$.
Montrer que $\forall n \in \mathbb{N}, U_{n+1} = \frac{1}{5}U_n A$, où A est la matrice de la partie A.
4. Exprimer U_n en fonction de A , n et U_0 .
5. En déduire a_n, b_n et c_n en fonction de n . Trouver les limites a, b et c des suites $(a_n), (b_n)$ et (c_n) .

Partie C

On donne le script Python suivant :

```
import numpy as np
import numpy.random as rd
def X(n): x=rd.randint(1,4)
for k in range(n): r=rd.rand()
if r <4/5:
if x+2 <=3:
else:
x=x+2
if .... : x=x-1
return(x)
```

1. Compléter l'instruction manquante de la fonction $X(n)$ pour qu'elle simule la variable X_n .
2. On ajoute les lignes suivantes :

```
N=10000 n=1000
Ech = []
for k in range(N):
Ech=Ech+[X(n)] print(np.mean(Ech))
```

Quel devrait être sensiblement l'affichage obtenu ?

Exercice 12. [Oral HEC] Une coccinelle se déplace sur un tétraèdre régulier PQRS (une pyramide) en longeant les arêtes. Elle est placée à l'instant $n = 0$ sur le sommet P . On suppose que, si, elle se,

trouve sur un sommet à l'instant n , elle sera sur l'un des trois autres sommets à l'instant $n + 1$ de, façon équiprobable.

Pour tout $n \in \mathbb{N}$ on note p_n (respectivement q_n , r_n et s_n) la probabilité que la coccinelle se trouve sur le sommet P (respectivement Q , R et S) à l'instant n .

1. On définit la matrice ligne X_n par

$$X_n = (p_n \quad q_n \quad r_n \quad s_n)$$

Déterminer une matrice, carrée $A \in \mathcal{M}_4(\mathbb{R})$, indépendante de n , telle que $X_{n+1} = X_n A$ pour tout $n \in \mathbb{N}^*$

2. Montrer que cette matrice est diagonalisable.
3. Calculer A^n pour tout $n \in \mathbb{N}^*$; donner une expression de p_n , q_n , r_n et s_n en fonction de n .
4. On note T le temps nécessaire à la coccinelle pour visiter au moins 3 sommets distincts du tétraèdre. Déterminer la loi de T . Calculer son espérance.
5. On note U le temps nécessaire à la coccinelle pour visiter tous les sommets du tétraèdre. Montrer que pour tout entier ℓ supérieur ou égal à 3

$$P(U = \ell) = \frac{2^{\ell-1} - 2}{3^{\ell-1}}$$

6. Déterminer l'espérance de U si elle existe.

Exercice 13. [Oral HEC] On étudie la vente d'un certain type de produit sur internet sur trois sites A , B , C et on fait les constatations suivantes :

- si un client choisit le site A pour un achat, il choisit indifféremment A , B ou C pour l'achat suivant,
- si un client fait un achat auprès du site B , il fait l'achat suivant sur le même site B ,
- si un client fait un achat sur le site C , il choisira pour l'achat suivant le site A avec une probabilité $1/12$, le site B avec une probabilité $7/12$ et le site C avec une probabilité $1/3$.

Au départ le client choisit au hasard l'un des trois sites.

On suppose que l'expérience est modélisée par un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Pour $n \in \mathbb{N}^*$, on désigne par p_n , q_n et r_n les probabilités pour que, au n -ième achat, le client se fournisse respectivement auprès de A , B et C .

1. Quelles sont les valeurs de p_1 , q_1 et r_1 ?
2. Pour tout $n \in \mathbb{N}^*$, donner une relation entre p_n , q_n et r_n .
3. Exprimer p_{n+1} , respectivement q_{n+1} et r_{n+1} , en fonction des trois réels p_n , q_n et r_n .
4. Pour $n \geq 2$, exprimer p_n en fonction de r_n et r_{n-1} .
5. Prouver que la suite $(r_n)_{n \in \mathbb{N}^*}$ est une suite récurrente linéaire. Donner l'expression de r_n , puis p_n et q_n en fonction de n .
6. Étudier la convergence des trois suites $(r_n)_{n \in \mathbb{N}^*}$, $(p_n)_{n \in \mathbb{N}^*}$ et $(q_n)_{n \in \mathbb{N}^*}$.