Modélisation de l’incertitude (M2 MIAGE IA²)

Andrea G. B. Tettamanzi
Laboratoire I3S – Équipe SPARKS
andrea.tettamanzi@univ-cotedazur.fr

univ-cotedazur.fr

Séance 1
Introduction

Objectifs de cet enseignement

- Fournir les notions de base des principales approches de la modélisation de l'incertitude en Intelligence Artificielle
  - Probabilités, réseaux bayésiens
  - Modèles de Markov cachés
  - Théorie de l'évidence de Dempster-Shafer
  - Théorie des ensembles flous, logique floue
  - Théorie des possibilités

Dans cette séance

- L'incertitude en Intelligence Artificielle
- Rappels de théorie des probabilités

Introduction

- L’incertitude est partout dans l’information et dans les connaissances
- La prise en compte de l’incertitude dans les systèmes d’inférence a longtemps été un problème en IA
- Les formalismes logiques ont dominé l’IA pendant plusieurs décennies
  - Logiques modales
  - Logique non-monotone
  - Logiques multi-valeurs (p.ex., logique floue)
- Les réseaux bayésiens ont gagné d’importance et ont permis d’intégrer les probabilités en IA

Deux types d'incertitude

- Épistémique
  – Connaissances incomplètes, ignorance
  – Le vrai état du phénomène n'est pas connu précisément, même s'il est déterministe
- Ontique
  – L'incertitude est inhérente dans le phénomène
  – Le résultat d'une expérience (action) est imprévisible, mais suit des lois
  – Exemple : lancer d'un dé

Qualification

- Qualifier l'incertitude signifie énumérer toutes les possibilités, même celles les moins plausibles
- Exemple : diagnose d'une maladie
  - On observe un symptôme (fièvre)
  - Plusieurs maladies pourraient en être la cause : grippe, coup de soleil, paludisme, chikungunya, covid-19, …
  - L'apparition de ce symptôme n'est même pas certaine pour ces maladies !

Représentation de croyances

Dès lors que la connaissance est incomplète, nous devons parler de **croyances**

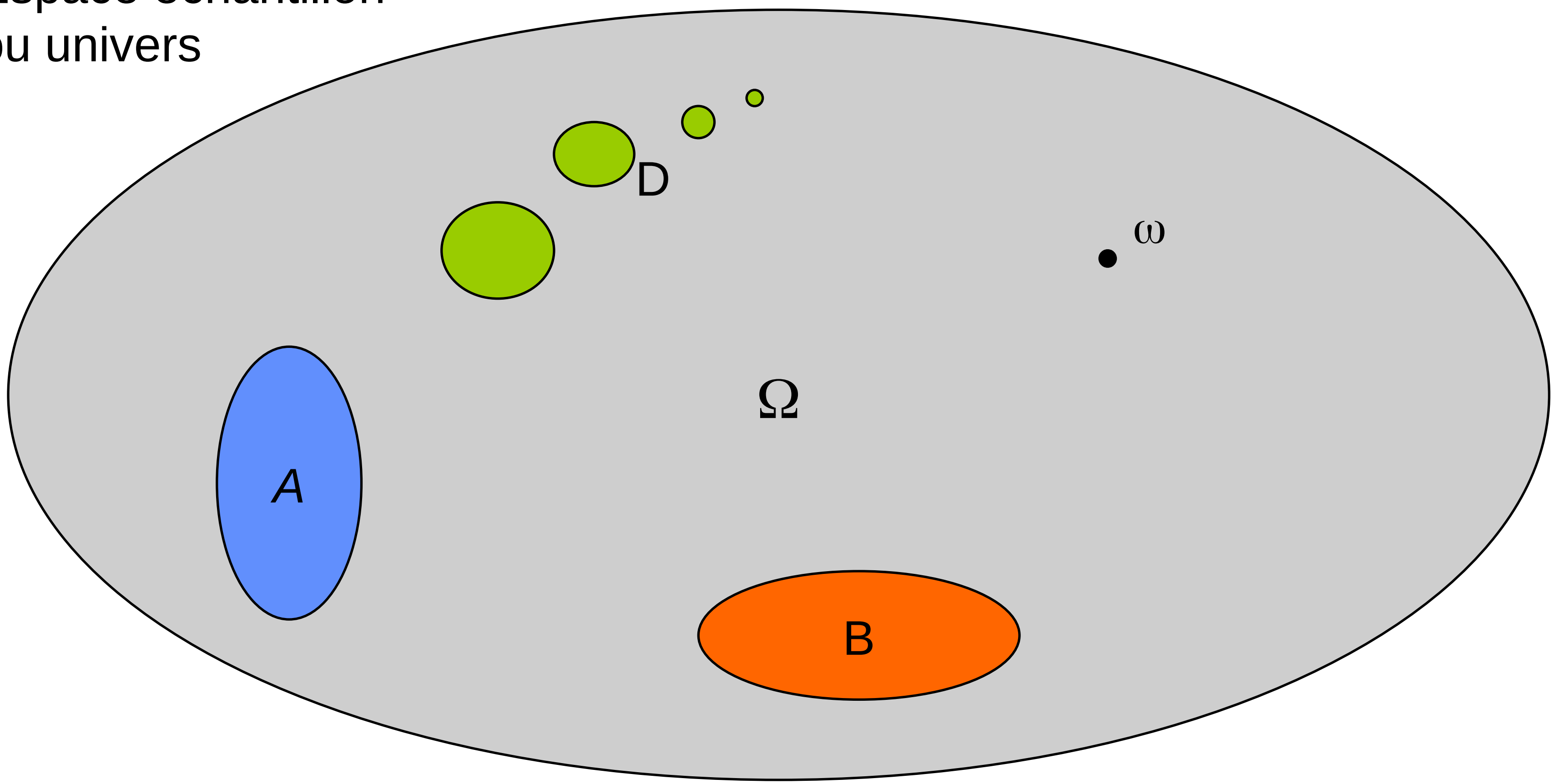
Un agent intelligent doit maintenir un état des croyances sur le monde qui l'entoure.

Il semble y avoir trois traditions de la représentation des croyances

- Fonctions sur des ensembles
- Logiques multi-valeurs
- Logiques modales

Événements

Espace échantillon
ou univers



Fonctions sur des ensembles

- Une fonction sur des ensembles est utilisée pour affecter des degrés de croyance à des propositions
- Une proposition est représentée par l'ensemble de ses modèles

Mais Spohn:

$$f : 2^\Omega \to [0, 1] \qquad \kappa : 2^\Omega \to \mathbb{N}$$

$$A \subseteq B \Leftrightarrow f(A) \leq f(B)$$

$$f(\emptyset) = 0$$

$$f(\Omega) = 1$$

Probabilité
Possibilité et nécessité
Croyance et plausibilité
Probabilités imprécises

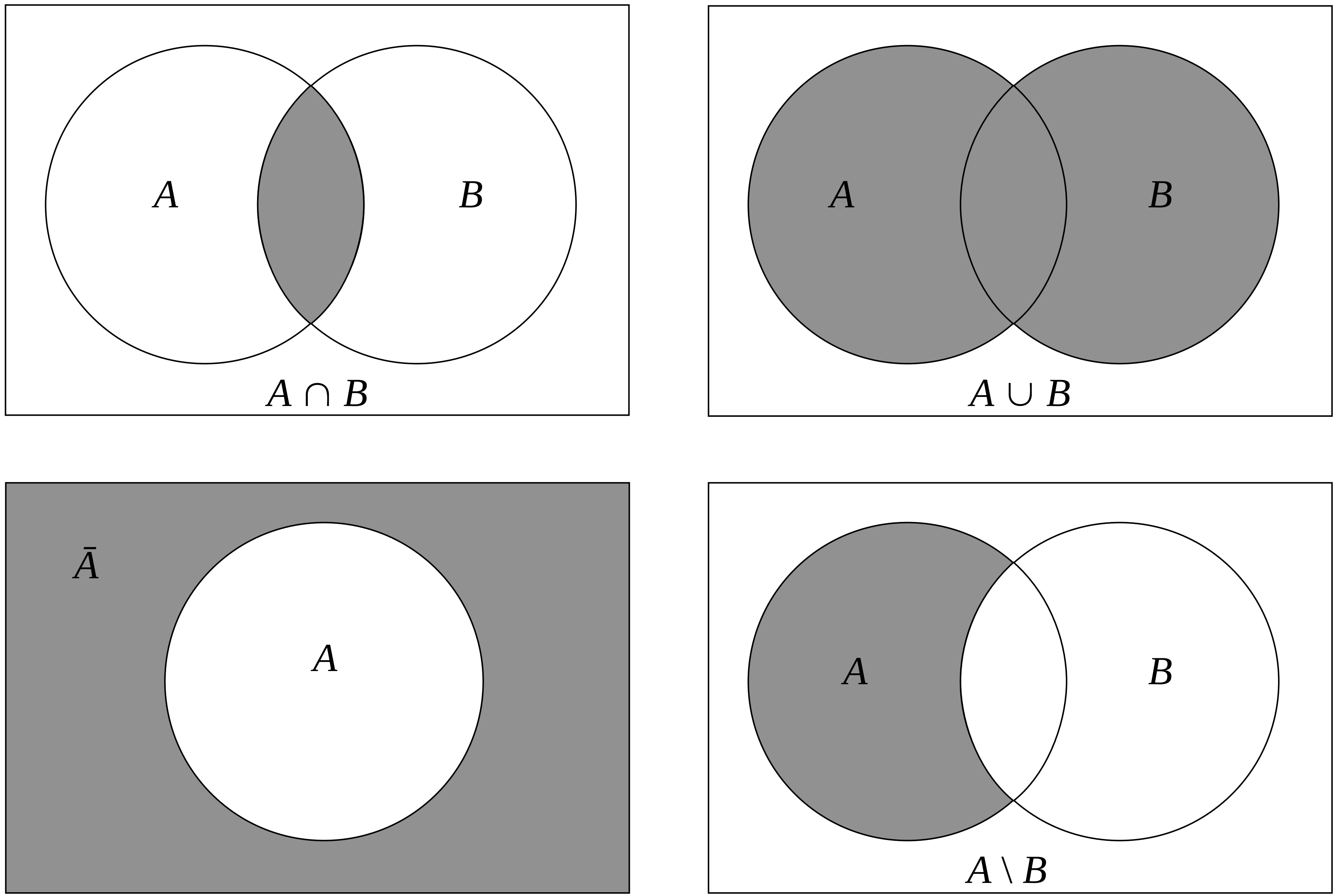
Probabilités

$$p : \Omega \to [0, 1]$$

$$\sum_{\omega \in \Omega} p(\omega) = 1$$

$$\Pr : 2^{\Omega} \to [0, 1]$$

$$\Pr(\phi) = \sum_{\omega \models \phi} p(\omega)$$


A
B
A ∩ B
A
B
A ∪ B
Ā
A
A
B
A \ B

Quelques propriétés

$$1 = P(\Omega) = P(\Omega + \emptyset) = P(\Omega) + P(\emptyset) = 1 + P(\emptyset)$$

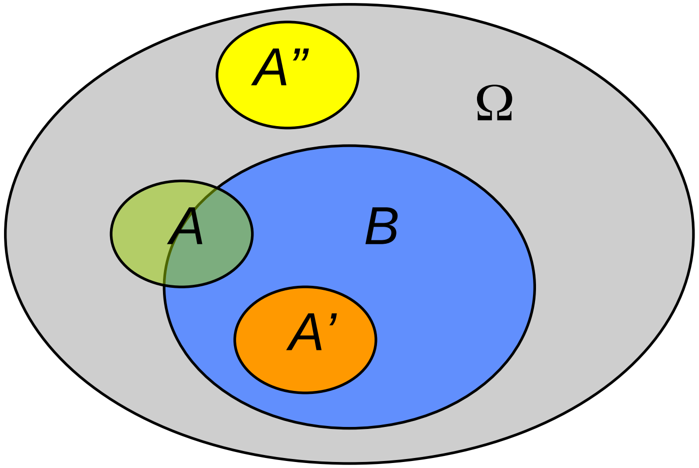
$$\therefore P(\emptyset) = 0$$

$$P(A) + P(\overline{A}) = P(A \cup \overline{A}) = P(\Omega) = 1$$

$$\therefore P(\overline{A}) = 1 - P(A)$$

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$$

Probabilité conditionnelle



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

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

$$P(A'' \mid B) = 0$$

Probabilité de l’intersection de deux événements

$$P(A \cap B) = P(B)P(A \mid B) = P(A)P(B \mid A)$$

Indépendence statistique

Deux événements A et B sont ***indépendents*** si et seulement si

$$P(A \mid B) = P(A)$$

Si A et B sont indépendents,

$$P(A \cap B) = P(A)P(B)$$

Formule de Bayes

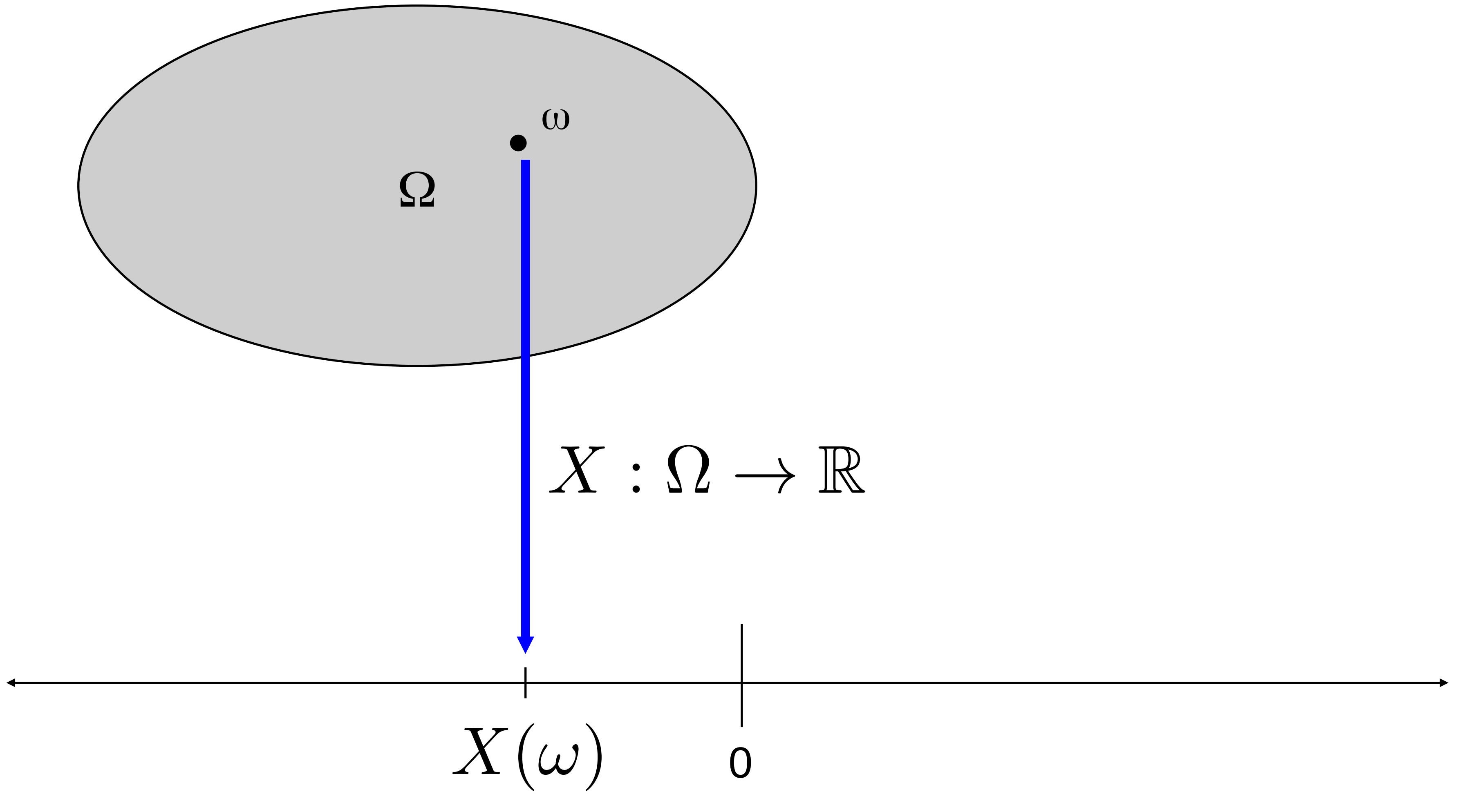
$$\Omega = A_1 \cup A_2 \cup \ldots \cup A_n$$ mutuellement disjoints

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(A_i)P(B \mid A_i)$$

$$P(A_i \mid B)P(B) = P(A_i \cap B) = P(A_i)P(B \mid A_i)$$

$$\boxed{P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(B)} = \frac{P(A_i)P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j)P(B \mid A_j)}}$$

Variables aléatoires



Fonction de distribution de probabilité

Appelée aussi :

fonction de repartition

fonction cumulative de probabilité

$$F(x) = P(X \le x)$$

1

$$\lim_{x \to -\infty} F(x) = 0 \qquad \lim_{x \to +\infty} F(x) = 1$$

2

$$x_1 < x_2 \Rightarrow F(x_1) \le F(x_2)$$

3

$$\lim_{h \to 0^+} F(x+h) = F(x)$$

Fonction de probabilité discrète

$$f(x) = \begin{cases} P(X = x_i) & x = x_i \\ 0 & x \neq x_i \end{cases}$$

$$\sum_i f(x_i) = 1$$

Fonction de densité de probabilité

$$f(x) = \frac{dF(x)}{dx}$$

$$P(x_0 < X \leq x_1) = \int_{x_0}^{x_1} f(x)dx$$

Espérance mathématique

Variable aléatoire discrète :

$$E[X] = \sum_i x_i f(x_i) = \mu$$

Variable aléatoire continue :

$$E[X] = \int_{-\infty}^{+\infty} x f(x) dx$$

Fonction de variable aléatoire :

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

Variance

Variable aléatoire discrète :

$$\mathrm{var}[X] = \sum_i (x_i - E[X])^2 f(x_i) = \sigma^2$$

Variable aléatoire continue :

$$\mathrm{var}[X] = \int_{-\infty}^{+\infty} (x - E[X])^2 f(x) dx$$

Proprieté : $\mathrm{var}[X] = E[X^2] - E^2[X]$

Écart type : $\sigma_X = \sqrt{\mathrm{var}[X]}$

Processus Stochastiques

Une suite de variables aléatoires

$$X_1, X_2, \ldots, X_t, \ldots$$

Chacune dotée de sa propre distribution de probabilité.

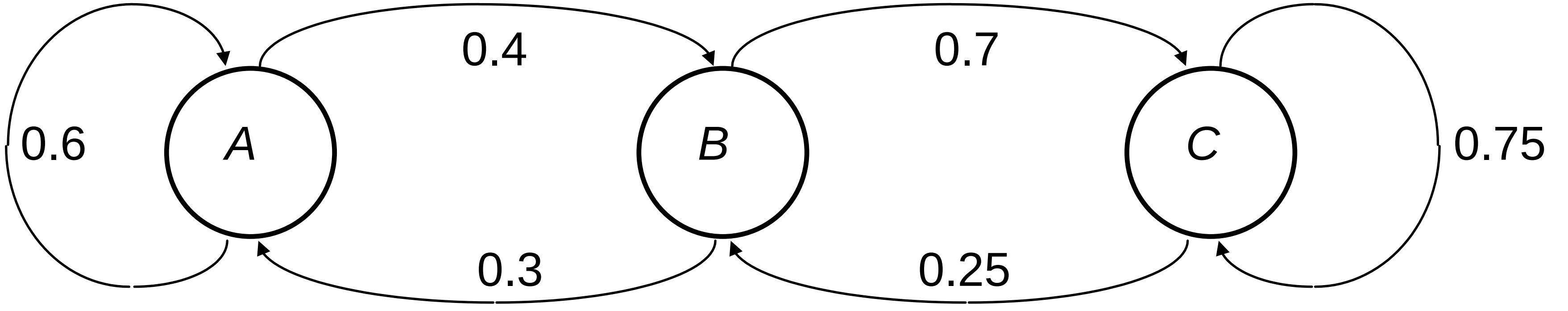
Notation: $$\{X_t(\omega)\}_{t=0,1,\ldots}$$

Chaînes de Markov

Un processus stochastique $\{X_t(\omega)\}_{t=0,1,\ldots}$

est une chaîne de Markov si et seulement si, pour tout t,

$$\Pr[X_t = x \mid X_0, X_1, \ldots, X_{t-1}] = \Pr[X_t = x \mid X_{t-1}]$$



Matrice de transition

$$\mathbf{T} = \begin{bmatrix} \Pr(X_t = x_1 \mid X_{t-1} = x_1) & \dots & \Pr(X_t = x_n \mid X_{t-1} = x_1) \\ \Pr(X_t = x_1 \mid X_{t-1} = x_2) & \dots & \Pr(X_t = x_n \mid X_{t-1} = x_2) \\ \vdots & & \vdots \\ \Pr(X_t = x_1 \mid X_{t-1} = x_n) & \dots & \Pr(X_t = x_n \mid X_{t-1} = x_n) \end{bmatrix}$$

T est une matrice stochastique :

$$\forall i, \quad \sum_{j=1}^{n} \Pr(X_t = x_j \mid X_{t-1} = x_i) = 1$$