Modélisation de l'incertitude (M2 MIAGE IA²)

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

univ-cotedazur.fr

Séance 5
Théorie de l'évidence

Dans cette séance

- Introduction
- Deux paradoxes
- Théorie de Dempster-Shafer
- Modèle des croyances transférables

Introduction

- Dans les sciences et l’ingénierie, nous devons toujours raisonner en présence de
  - Connaissances partielles
  - Informations incertaines
- On a déjà parlé des deux aspects de l’incertitude :
  - Ontique, ou aléatoire, dépendant de la variabilité des résultats d’expériences répétables
  - Épistémique, dépendant d’une connaissance insuffisante ou de l’impossibilité de distinguer/observer certaines caractéristiques d’un phénomène
- La théorie des probabilité n’est pas toujours adaptée pour traiter cette dernière, il faut donc d’autres théories plus adaptées

Paradoxe de l’eau et du vin

- Le principe d’indifférence (PI) en probabilités dit qu’en l’absence d’informations sur une quantité X, il faut assigner une probabilité égale a chacune des valeurs possibles de X
- Supposons qu’on nous donne un mélange d’eau et de vin. Tout ce qu’on sait est que le rapport vin/eau, X, est compris entre 1/3 et 3.
- Quelle est $P(X \leq 2)$? On utilise le PI :
  - $X \sim U[1/3, 3] \Rightarrow P(X \leq 2) = (2 - 1/3)/(3 - 1/3) = 5/8$
- Maintenant, soit $Y = 1/X$ (rapport eau/vin) ; on aura
  - $Y \sim U[1/3, 3] \Rightarrow P(Y \geq 1/2) = (3 - 1/2)/(3 - 1/3) = 15/16$
- Mais $Y \geq 1/2$ si et seulement si $X \leq 2$, donc $P(Y \geq 1/2) = P(X \leq 2)$!
- Donc, $P(X \leq 2) = 5/8$, mais, **en même temps**, $P(X \leq 2) = 15/16$!?

Paradoxe d'Ellsberg

- On a une urne contenant 30 balles rouges et 60 balles noires ou jaunes
- On a le choix entre deux paris :
  - A : on gagne 100 € si on tire une balle rouge
  - B : on gagne 100 € si on tire une balle noire
- On a aussi le choix entre ceux deux autres paris :
  - C : on gagne 100 € si on tire une balle rouge ou jaune
  - D : on gagne 100 € si on tire une balle noire ou jaune
- Or, la plupart des gens préfère strictement A à B
  - On devrait donc en conclure que **P(rouge) > P(noire)**
- Mais ils préfèrent aussi strictement D à C
  - On devrait donc en conclure que **P(noire) > P(rouge)**

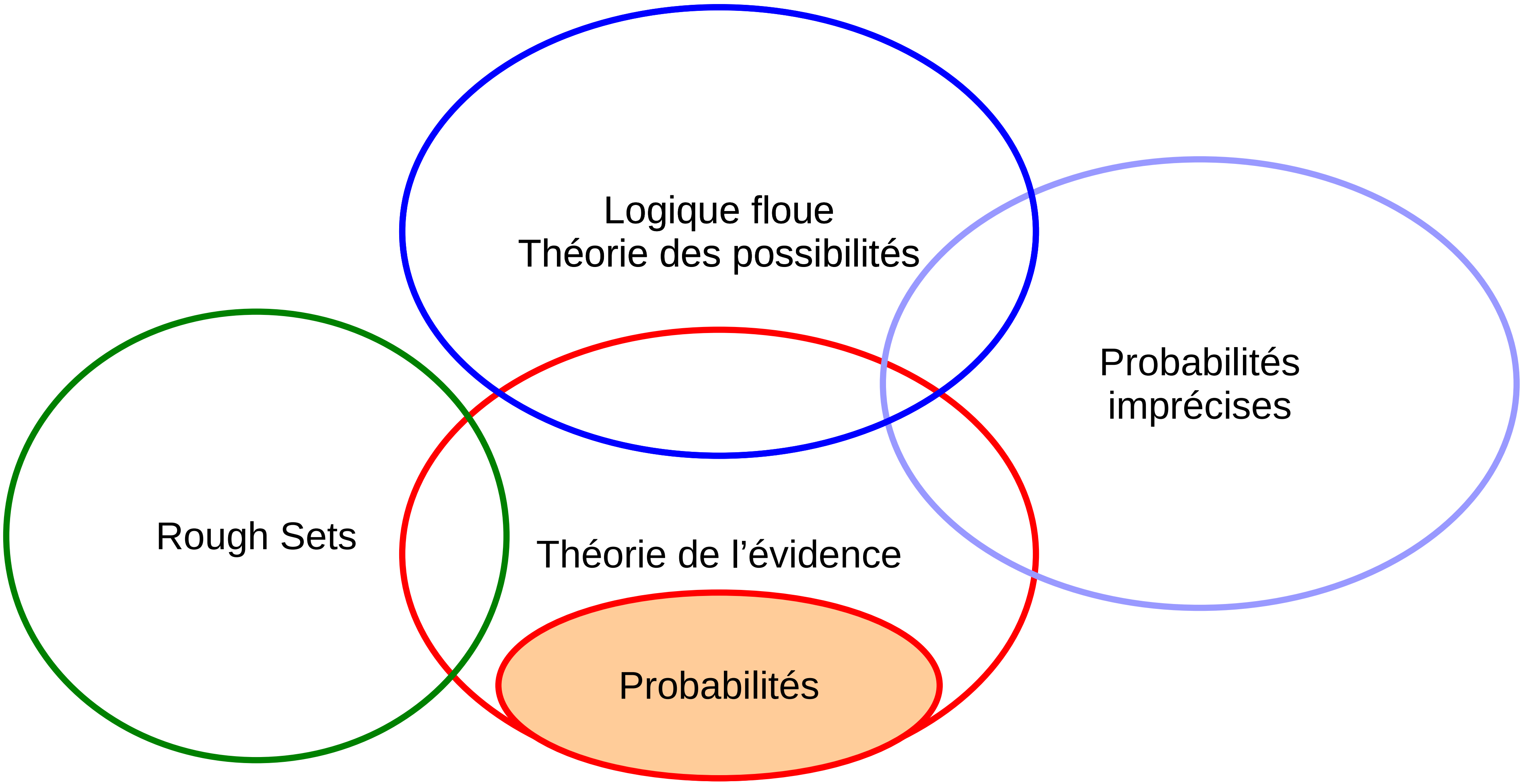
Théorie de Dempster-Shafer

- Cadre formel pour représenter et raisonner avec des information incertaines
- Connue aussi sous le nom de théorie de l’évidence ou théorie des fonctions de croyance
- Originée par les travaux d’Arthur Dempster (1968) sur l’inférence statistique
- Formalisée par Glenn Shafer (1976) comme théorie de l’évidence
- Ultérieurement développée par Philippe Smets dans les 1980 et 1990 sous le nom de théorie des croyances transférables.

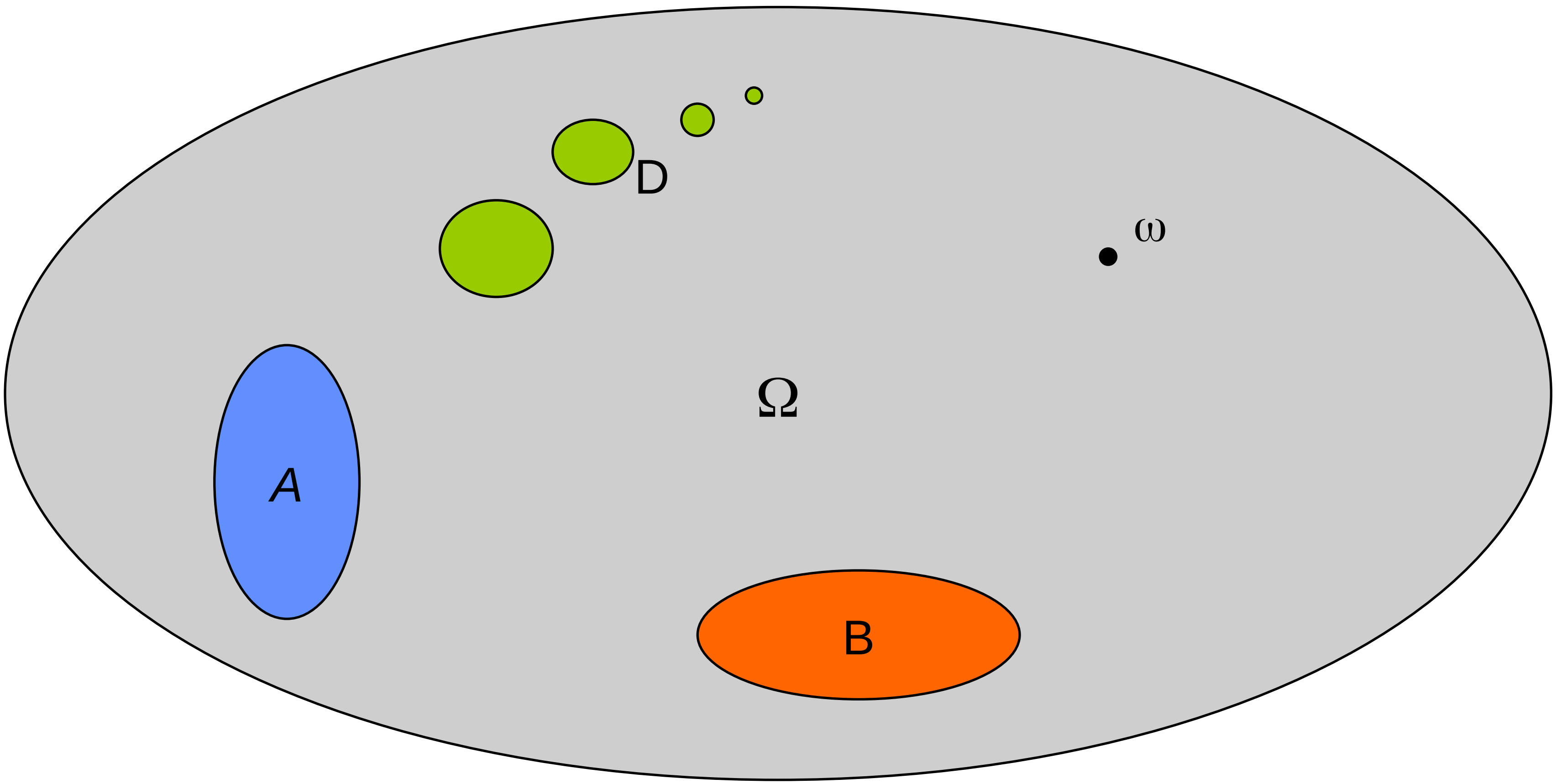
Idée de fond

- La théorie de l'évidence étend la théorie des probabilités
- Une fonction de croyance peut être considérée à la fois comme un ensemble généralisé et comme une mesure non additive
- La théorie comprend des extensions de notions probabilistes (conditionnement, marginalisation) et des notions ensemblistes (intersection, union, inclusion, etc.)
- Le raisonnement de Dempter-Shafer donne les mêmes résultats que le raisonnement probabiliste lorsqu'on lui fournit les mêmes informations
- Cependant, sa plus grande expressivité nous permet de représenter ce que nous savons de manière plus fidèle

Relations avec d'autres théories



Cadre de discernement



Fonction de masse

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

$$\sum_{A \subseteq \Omega} m(A) = 1$$

$A \subseteq \Omega : m(A) > 0$ est un **ensemble focal**

Fonction de masse normalisée $\Leftrightarrow$ $m(\emptyset) = 0$

Mesures de croyance et plausibilité

$$bel(B) = \sum_{A \subseteq B} m(A)$$

$$pl(B) = \sum_{A \cap B \neq \emptyset} m(A) = 1 - bel(\overline{B})$$

$$bel(B) \leq pl(B)$$

bel(*B*) = *pl*(*B*) lorsque les ensembles focaux de *m* sont des singletons (événements élémentaires) ; *m* est alors dite bayésienne et *bel*() est une mesure de probabilité.

Combinaison d'évidence : la règle de Dempster

$$(m_1 \oplus m_2)(\emptyset) = 0$$

$$(m_1 \oplus m_2)(A) = \frac{1}{1-K} \sum_{B \cap C = A \neq \emptyset} m_1(B) m_2(C)$$

$$K = \sum_{B \cap C = \emptyset} m_1(B) m_2(C)$$ niveau de conflit entre les deux masses

C'est une généralisation de la règle de Bayes !
Si on combine *m* avec une masse toute concentrée sur *A*, on obtien le conditionnement de Dempster, *m*(. | *A*)

Ensemble crédal

- Une mesure de probabilité *P* est compatible avec une masse *m* si

$$\forall A \subseteq \Omega, \quad bel(A) \leq P(A) \leq pl(A)$$

- L'ensemble de toutes les mesures de probabilités compatible avec *m* est appelé l'ensemble crédal de *m*

$$\mathcal{P}(m) = \{P : \forall A \subseteq \Omega, bel(A) \leq P(A)\}$$

- La mesure *bel*() est son enveloppe inférieure

Modèle des croyances transférables

- Modèle non probabiliste de raisonnement incertain reposant sur la théorie de l’évidence
- Proposé et développé par Philippe Smets au début des années 1990
- Le modèle des croyances transférable comporte deux niveaux :
  - un niveau crédal pour représenter et combiner les informations,
  - un niveau pignistique pour prendre une décision

Exemple de Zadeh

- Un patient a une maladie qui peut être causée par trois facteurs différents A, B ou C.
  - Le médecin 1 dit que la cause est très probablement A (p = 0,95), mais B est également possible mais peu probable (p = 0,05).
  - Le médecin 2 dit que la cause est très probablement C (p = 0,95), mais que B est également possible mais peu probable (p = 0,05).
- Que doit-on en penser ?
  - L'actualisation bayésienne de la première opinion par la seconde (ou l'inverse) implique la certitude que la cause est B.
  - La règle de combinaison de Dempster conduit au même résultat.
- Cela peut être considéré comme paradoxal, car bien que les deux médecins pointent des causes différentes, A et C, ils sont tous les deux d'accord pour dire que B n'est pas probable !

Idée de fond du modèle des croyances transférables

- Contrairement à Dempster-Shafer, Smets fait l'**hypothèse du monde ouvert**, qui relaxe l'hypothèse selon laquelle tous les résultats possibles sont connus
- Dans l'hypothèse du monde ouvert, la règle de combinaison de Dempster est adaptée de telle sorte qu'il n'y a **pas de normalisation**
- La masse de l'ensemble vide est prise pour indiquer un résultat inattendu, par exemple la croyance en une hypothèse **hors du cadre du discernement**
- Cette adaptation viole le caractère probabiliste de la théorie originale et également l'inférence bayésienne
- Par conséquent, on substitue
  - masses de probabilité → degrés de croyance
  - mise à jour des probabilités → transfert
- D’où le nom du modèle

Règle de combinaison

$$(m_1 \oplus m_2)(A) = \sum_{B \cap C = A} m_1(B) m_2(C)$$

$$A, B, C \neq \emptyset$$

Niveau pignistique

$$P_{\mathrm{Bet}}(\omega) = \sum_{\omega \in A} \frac{m(A)}{|A|}$$

|*A*| dénote le nombre d'événements élémentaires (singletons) qui composent A.

Probabilité **pignistique** : en théorie des décisions, la probabilité qu'un être rationnel affecte à une option lors qu'il doit prendre une décision.

Cette définition récupère le principe d'indifférence !