

MODE — cours 3:  
fonctions de croyance et décisions  
séquentielles

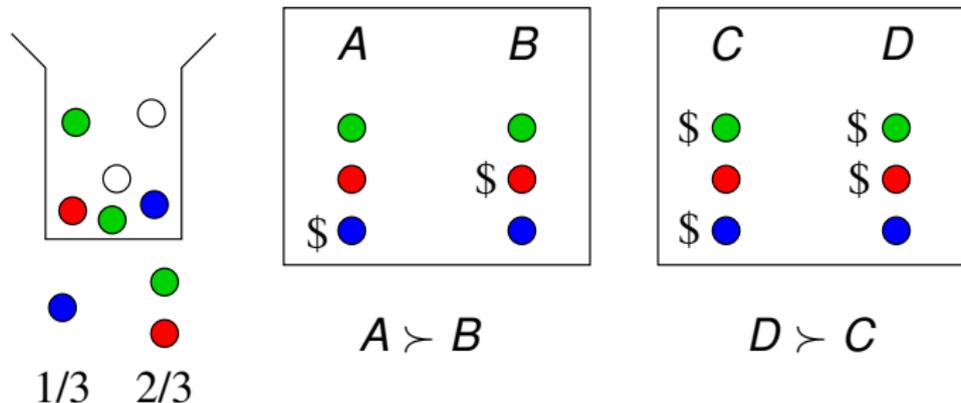
Christophe Gonzales

LIP6 – Université Paris 6, France

- 1 Fonctions de croyance
- 2 Modèles décisionnels avec fonctions de croyance
- 3 Décisions séquentielles avec EU
- 4 Décisions séquentielles avec RDU

# 1 Les fonctions de croyance

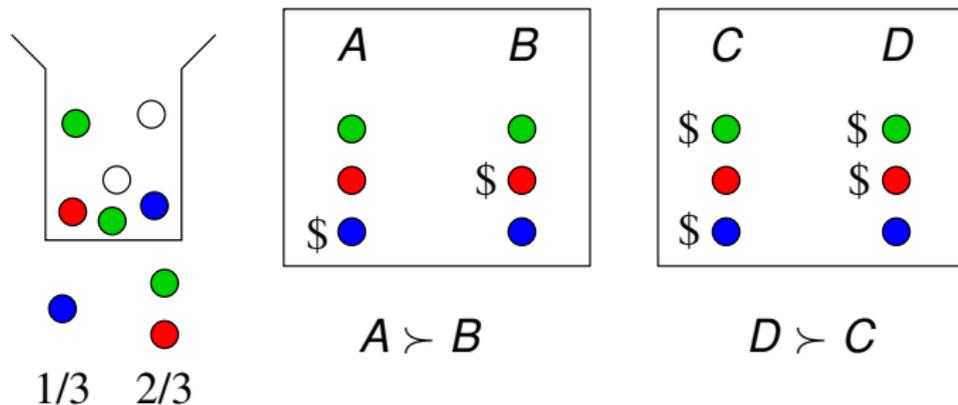
# Retour à l'urne d'Ellsberg



Violation du Sure thing principle

Peut-on représenter  $\succ$  par RDU ?

# L'urne d'Ellsberg et RDU



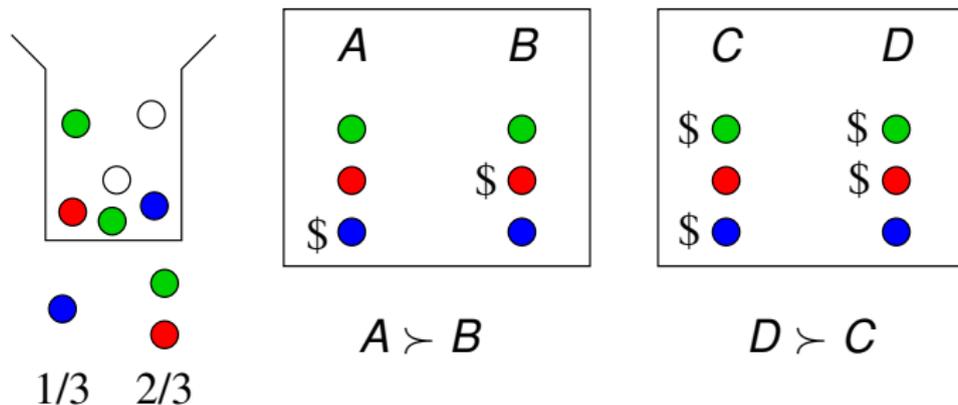
Probas  $P_r, P_b, P_v$

$$A \succ B \iff \varphi(P_b) > \varphi(P_r)$$

$$D \succ C \iff \varphi(P_r + P_v) > \varphi(P_b + P_v)$$

or  $\varphi$  croissante  $\implies$  RDU impossible

# L'urne d'Ellsberg et les capacités



Problème :  $\varphi(P_b) > \varphi(P_r)$  et  $\varphi(P_r + P_v) > \varphi(P_b + P_v)$

## Définition d'une capacité

- fonction  $\mu$  des sous-ensembles de  $\mathcal{S}$  dans  $[0, 1]$
- $\mu(\emptyset) = 0$ ,  $\mu(\mathcal{S}) = 1$
- monotone :  $\forall A \subset B \subseteq \mathcal{S}, \mu(A) \leq \mu(B)$

L'urne d'Ellsberg  $\implies$  capacités à la Choquet

capacités  $\xrightarrow{?}$  représentation des incertitudes

Théorie des fonctions de croyance

## Fonction de croyance (Dempster-Shafer)

- $\mathcal{S}$  : ensemble des états de la nature
- $\mathcal{A}$  : ensemble des sous-ensembles de  $\mathcal{S}$
- $\mu : \mathcal{A} \mapsto [0, 1]$  fonction de croyance :
  - $\mu(\emptyset) = 0, \quad \mu(\mathcal{S}) = 1$
  - $\mu$  monotone d'ordre  $\infty : \forall n \geq 2, \forall A_1, \dots, A_n \in \mathcal{A} :$

$$\mu \left( \bigcup_{i=1}^n A_i \right) \geq \sum_{\emptyset \subset I \subseteq \{1, \dots, n\}} (-1)^{|I|+1} \mu \left( \bigcap_{i \in I} A_i \right)$$

Ordre 1 :  $\forall A \subseteq B \subseteq \mathcal{S}, \mu(A) \leq \mu(B)$

$\implies$  capacités à la Choquet

Ordre 2 :  $\mu(A \cup B) \geq \mu(A) + \mu(B) - \mu(A \cap B)$

# L'inverse de Möbius

## *Définition de l'inverse de Möbius*

Fonction  $\phi : \mathcal{A} \mapsto [0, 1]$  définie par :

$$\phi(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \mu(B)$$

## *Expression de la fonction de croyance*

$$\mu(A) = \sum_{B \subseteq A} \phi(B)$$

## *Proposition*

$\mu$  définie à partir de  $\phi$  est une fonction de croyance ssi :

- $\phi(\emptyset) = 0$
- $\sum_{B \subseteq \mathcal{S}} \phi(B) = 1$
- $\forall B \subseteq \mathcal{S}, \phi(B) \geq 0$

## Exemple : le camionneur

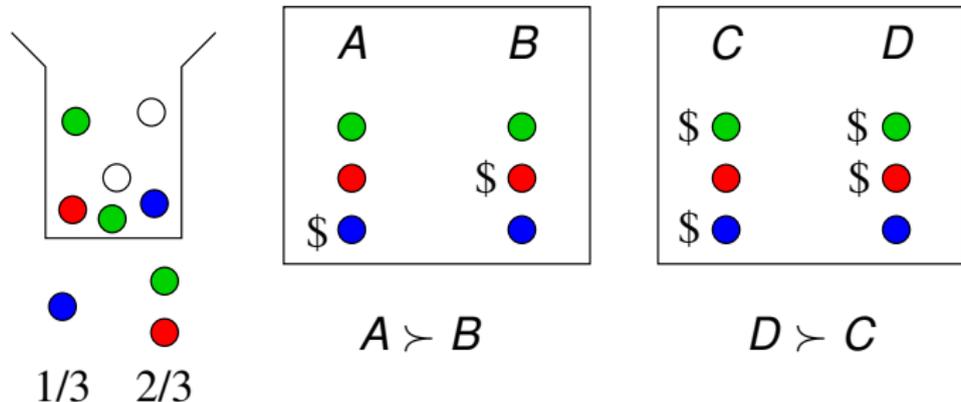
pièces de rechange :  $\begin{cases} \text{G(ood)} \\ \text{A(cceptable)} \\ \text{B(ad)} \end{cases}$

**Information** : au plus la moitié des pièces est d'un type donné

A	$\emptyset$	{G}	{A}	{B}	{G, A}	{G, B}	{A, B}	$\mathcal{S}$
$\mu$	0	0	0	0	1/2	1/2	1/2	1
$\phi$	0	0	0	0	1/2	1/2	1/2	-1/2

$\implies \mu$  n'est pas une fonction de croyance

# Exemple : l'urne d'Ellsberg



<i>Evt</i>	$\emptyset$	$\{B\}$	$\{R\}$	$\{V\}$	$\{B, R\}$	$\{B, V\}$	$\{R, V\}$	$\mathcal{S}$
$\mu$	0	$1/3$	0	0	$1/3$	$1/3$	$2/3$	1
$\phi$	0	$1/3$	0	0	0	0	$2/3$	0

- $\phi \implies$  masses de Möbius
- $\phi(B) > 0$  : éléments focaux
- éléments focaux = singletons  $\iff \mu = \text{proba}$
- fonction de croyance = enveloppe inférieure de lois de proba

## ② Modèles décisionnels avec fonctions de croyance

# Décisions avec des fonctions de croyance

- $\mathcal{S}$  : ensemble des états de la nature
- $\mathcal{A}$  : ensemble des sous-ensembles de  $\mathcal{S}$
- $f : \mathcal{A} \mapsto [0, 1]$  fonction de croyance (d'inverse de Möbius  $\phi$ )
- $\mathcal{X}$  : ensemble des conséquences
- $\mathcal{Y}$  : ensemble des sous-ensembles de  $\mathcal{X}$
- décision  $\delta \equiv$  acte = fonction  $\mathcal{S} \mapsto \mathcal{X}$   
 $\implies \forall Y \in \mathcal{Y}, g(Y) = f(\delta^{-1}(Y)) =$  incertitude sur l'obtention de  
 $Y$  en prenant la décision  $\delta$

$g$  est une fonction de croyance

$\psi : \mathcal{Y} \mapsto [0, 1]$  définie par  $\psi(Y) = \sum_{\delta(B)=Y} \phi(B)$  est l'inverse de Möbius de  $g$

## Généralisation de vNM aux fonctions de croyance

$\mathcal{F}$  = ensemble de toutes les fonctions de croyance sur  $\mathcal{Y}$

### *Mixage de fonctions de croyance*

$\forall f, g \in \mathcal{F}, \forall \lambda \in [0, 1], h = \lambda f + (1 - \lambda)g \in \mathcal{F}$  :

fonction de croyance

$\forall Y \in \mathcal{Y}, h(Y) = \lambda f(Y) + (1 - \lambda)g(Y)$

$\implies$  comme dans von Neumann-Morgenstern, on va exprimer  $\succsim$  sur l'espace des mixages de fonctions de croyance  $\mathcal{F}$

## *Axiome 1 : Préordre total*

$\succsim$  est un préordre large total sur  $\mathcal{F}$

## *Axiome 2 : continuité*

$\forall f, g, h \in \mathcal{F}$  tels que  $f \succ g \succ h$ ,  $\exists \alpha, \beta \in ]0, 1[$  tels que :

$$\alpha f + (1 - \alpha)h \succ g \succ \beta f + (1 - \beta)h.$$

## *Axiome 2 : Indépendance*

$\forall f, g, h \in \mathcal{F}, \forall \alpha \in ]0, 1[$  :

$$f \succ g \implies \alpha f + (1 - \alpha)h \succ \alpha g + (1 - \alpha)h.$$

## *Théorème (von Neumann-Morgenstern)*

Les deux assertions suivantes sont équivalentes :

- 1  $\succsim$  sur  $\mathcal{F}$  vérifie les axiomes 1, 2, 3
- 2  $\succsim$  est représentable par  $U : \mathcal{F} \mapsto \mathbb{R}$  telle que :
  - $\forall f, g \in \mathcal{F}, f \succsim g \iff U(f) \geq U(g)$
  - $U$  est linéaire sur  $\mathcal{F}$ , i.e.,  $\forall f, g \in \mathcal{F}, \forall \lambda \in [0, 1]$  :

$$U(\lambda f + (1 - \lambda)g) = \lambda U(f) + (1 - \lambda)U(g)$$

De plus,  $U$  est unique à une transformation affine strictement positive près

Par récurrence :  $\forall \lambda_i \geq 0$  t.q.  $\sum_{i=1}^n \lambda_i = 1$ ,  $U\left(\sum_{i=1}^n \lambda_i f_i\right) = \sum_{i=1}^n \lambda_i U(f_i)$

*Fonction de croyance élémentaire concentrée sur B*

fonction de croyance  $e_B$  telle que :

$$e_B(A) = 1 \text{ si } A \supseteq B \quad \text{et} \quad e_B(A) = 0 \text{ sinon}$$

$\implies$  inverse de Möbius  $\psi_B$  telle que  $\psi_B(B) = 1$

*Ensemble focal*

- $f$  fonction de croyance d'inverse de Möbius  $\psi$
- $C_f =$  ensemble focal de  $f = \{B : \psi(B) > 0\}$

$\implies \forall$  fonction de croyance  $f$  et  $\forall$  événement  $A$  :

$$f(A) = \sum_{B \subseteq A} \psi(B) = \sum_{B \in C_f} \psi(B) e_B(A)$$

$$f = \sum_{B \in C_f} \psi(B) e_B$$

# Généralisation de EU (5/5)

$$\text{Rappel : } U \left( \sum_{i=1}^n \lambda_i f_i \right) = \sum_{i=1}^n \lambda_i U(f_i)$$

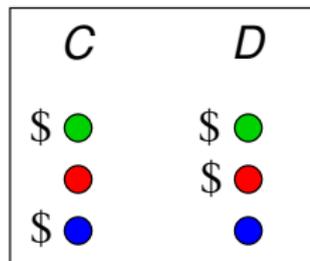
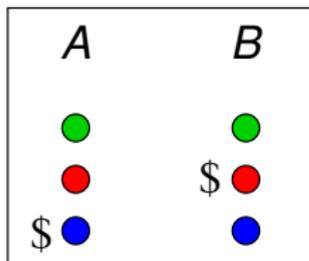
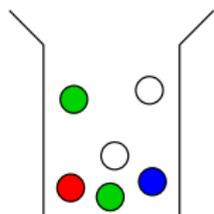
$$\text{Or } f = \sum_{B \in \mathcal{C}_f} \psi(B) e_B \implies U(f) = \sum_{B \in \mathcal{C}_f} \psi(B) U(e_B)$$

## *Belief expected utility*

- $f$  : fonction de croyance
- $\psi$  : inverse de Möbius
- $BEU(f) = \sum_{B \in \mathcal{C}_f} \psi(B) w(B)$

où  $w(B) = U(e_B)$  = utilité de l'ensemble de conséquences  $B$   
= utilité d'une croyance  $e_B$  dont le seul élément focal est  $B$

# BEU et l'urne d'Ellsberg



Evts	0	\$	{0, \$}
boules	{R, V}	{B}	S
$\mu$	2/3	1/3	1
$\phi$	2/3	1/3	0

$$BEU(A) = 2/3w(\{0\}) + 1/3w(\{\$\}) = 1/3$$

Evts	0	\$	{0, \$}
boules	{B, V}	{R}	S
$\mu$	1/3	0	1
$\phi$	1/3	0	2/3

$$BEU(B) = 1/3w(\{0\}) + 2/3w(\{0, \$\}) = 2/3\alpha$$

Evts	0	\$	{0, \$}
boules	{R}	{B, V}	S
$\mu$	0	1/3	1
$\phi$	0	1/3	2/3

$$BEU(C) = 1/3w(\{\$\}) + 2/3w(\{0, \$\})$$

Evts	0	\$	{0, \$}
boules	{B}	{R, V}	S
$\mu$	1/3	2/3	1
$\phi$	1/3	2/3	0

$$BEU(D) = 1/3w(\{0\}) + 2/3w(\{\$\}) = 2/3$$

*Problème de BEU :*

$$BEU(f) = \sum_{B \in \mathcal{C}_f} \phi(B)w(B) \implies w \text{ doit être défini sur } \mathcal{C}_f \text{ (voire } \mathcal{Y} = 2^{\mathcal{X}})$$

en comparaison :  $EU(P) = \sum_{x \in \mathcal{X}} P(x)u(x) \implies u$  défini sur  $\mathcal{X}$

$\implies$  EU nécessite moins d'élicitation d'utilité

## Axiome 4 : Dominance

- $\forall B \in \mathcal{Y} : m_B =$  la pire des conséquences dans  $B$
- $\forall B \in \mathcal{Y} : M_B =$  la meilleure des conséquences dans  $B$
- $\forall B \in \mathcal{Y} : e_B$  fonction de croyance élémentaire concentrée en  $B$
- $\forall B, B' \in \mathcal{Y}$ , si  $m_B \succsim_{\mathcal{X}} m_{B'}$  et  $M_B \succsim_{\mathcal{X}} M_{B'}$  alors  $e_B \succsim e_{B'}$

Idée :

aucune information sur les sous-événements inclus dans  $B$   
 $\implies$  décider en fonction de la pire et de la meilleure conséquence

## *Théorème — Jaffray (89)*

Les deux assertions suivantes sont équivalentes :

- ①  $\succsim$  sur  $\mathcal{F}$  vérifie les axiomes 1, 2, 3, 4
- ②  $\succsim$  est représentable par  $U : \mathcal{F} \mapsto \mathbb{R}$  telle que :
  - $\forall f, g \in \mathcal{F}, f \succsim g \iff U(f) \geq U(g)$
  - $\exists v$  définie sur  $\mathcal{M} = \{(m, M) \in \mathcal{X} \times \mathcal{X} : M \succsim_{\mathcal{X}} m\}$  t.q. :

$$U(f) = \sum_{B \in \mathcal{C}_f} \psi(B) v(m_B, M_B)$$

- $U$  et  $v$  sont uniques à une transformation affine croissante commune près
- $v$  est une fonction non décroissante en  $m$  et en  $M$
- $u(x) = v(x, x)$  est une utilité de vNM

# Le critère d'Hurwicz

Modèle de Jaffray  $\implies v$  défini sur  $|\mathcal{X}|^2$  éléments

$v(m, M) \implies$  attitude  $\left\{ \begin{array}{l} \text{vis à vis du risque} \\ \text{vis à vis de l'ambiguïté} \end{array} \right.$

## *Critère d'Hurwicz*

$\forall (m, M) \in \mathcal{M}, \alpha(m, M) =$  critère local d'optimisme/pessimisme

$\alpha(m, M) =$  valeur de  $\alpha$  pour laquelle le décideur est indifférent entre recevoir :

- 1  $m$  avec la proba  $\alpha$  et  $M$  avec la proba  $1 - \alpha$
- 2 au moins  $m$  et au plus  $M$ , sans autre information

$$v(m, M) = \alpha(m, M)u(m) + [1 - \alpha(m, M)]u(M)$$

pessimisme :  $\alpha(m, M) = 1$       optimisme :  $\alpha(m, M) = 0$

### ③ Décisions séquentielles avec EU

## Loteries et espérance d'utilité

- $\mathcal{X}$  : ensemble des conséquences
- $\mathcal{S}$  : ensemble des états de la nature
- acte : fonction  $\mathcal{S} \mapsto \mathcal{X}$
- $\mathcal{V}$  : ensemble des applications de  $\mathcal{S}$  dans  $\mathcal{X}$
- $\succsim_{\mathcal{D}}$  : relation de préférence sur  $\mathcal{V}$
- $f \in \mathcal{V} \implies P_f$  sur  $(\mathcal{X}, 2^{\mathcal{X}})$
- $\mathcal{L}$  : loteries, ensemble des lois à support fini sur  $\mathcal{X}$   
 $f \implies P_f = \langle c_1, p_1; \dots; c_n, p_n \rangle$ , avec  $c_1 \succsim_{\mathcal{X}} c_2 \succsim_{\mathcal{X}} \dots \succsim_{\mathcal{X}} c_n$
- $\succsim_{\mathcal{D}} \implies \succsim$  sur  $\mathcal{L}$
- $\succsim$  est représentable par  $U$  telle que

$$U(P) = \sum_{i=1}^n p_i u(c_i),$$

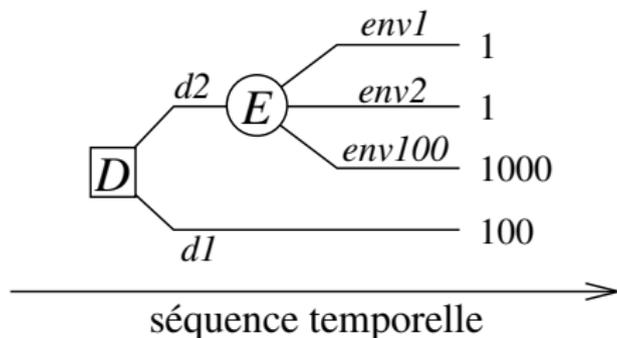
où  $u(c_i) = U(\langle c_i, 1 \rangle)$ .

# Les Arbres de décision

## Exemple

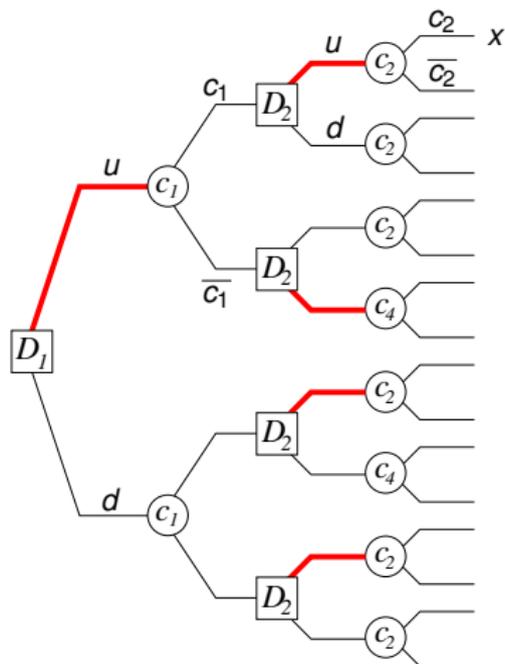
- enveloppe 1 contient 100 €
- enveloppe 2 choisie parmi une pile de 100 enveloppes dont 3 contiennent 1000 € et 97 contiennent 1 €

⇒  $\begin{cases} \text{enveloppe 1} = 100 \text{ €} \\ \text{enveloppe 2} = 3 \text{ chances sur } 100 \text{ d'avoir } 1000 \text{ € et} \\ \quad \quad \quad 97 \text{ chances sur } 100 \text{ d'avoir } 1 \text{ €} \end{cases}$



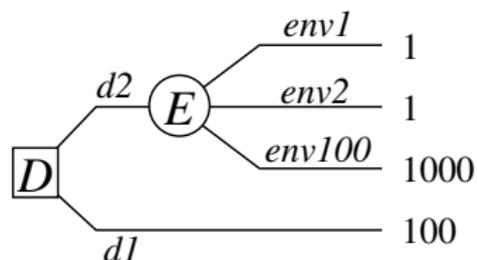
carrés = décisions    ronds = nœuds de chance

# Décisions séquentielles (1/3)



$x = \text{conséquence si } D_1 = u, C_1 = c_1, D_2 = u, C_2 = c_2$

## Décisions séquentielles (2/3)



⇒ la décision optimale est celle dont la moyenne des utilités des conséquences est la plus élevée

$d_1 \equiv$  loterie  $L_1 = \langle 100, 1 \rangle$

$d_2 \equiv$  loterie  $L_2 = \langle 1, 0.97; 1000, 0.03 \rangle$

$EU(d_1) = EU(L_1) = 100$

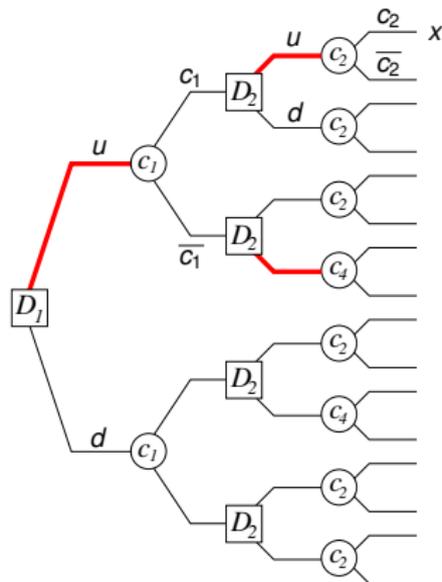
$EU(d_2) = EU(L_2) = 0,97 \times 1 + 0,03 \times 1000 = 30,97$

⇒ décision optimale selon EU :  $d_1$

# Décisions séquentielles (3/3)

## Stratégie

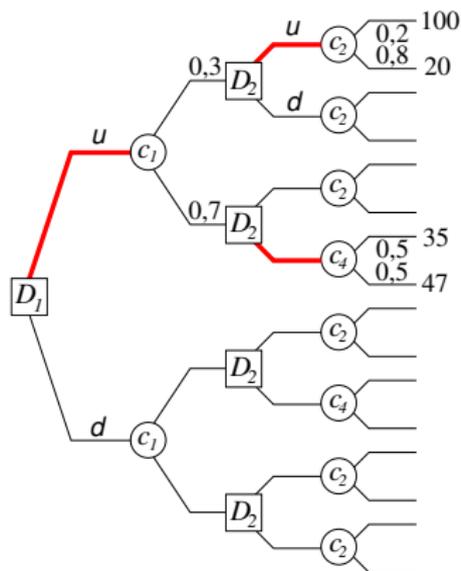
Une stratégie de décision = la sélection en **tout** sommet de décision  $D$  de l'arbre accessible compte tenu des décisions prises précédemment, d'une décision  $d$  appartenant à l'ensemble des décisions réalisables de ce sommet.



# Décisions séquentielles (3/3)

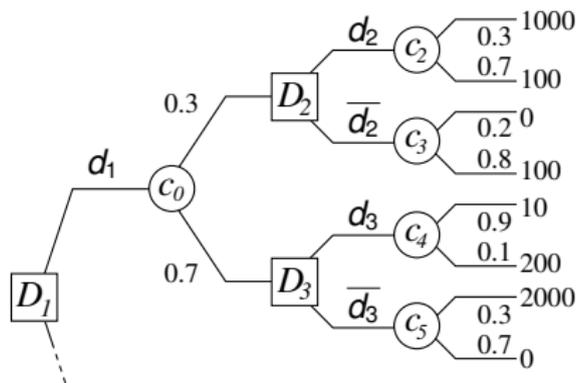
## Stratégie et espérance d'utilité

- À toute stratégie correspond une loterie
- critère d'optimalité = espérance max sur les loteries



loterie =  $\langle 20, 0.24; 35, 0.35; 47, 0.35; 100, 0.06 \rangle$

# Calculs dans un arbre de décision (1/5)



Stratégie  $S_1 = \langle\langle D_1 = d_1, D_2 = d_2, D_3 = d_3 \rangle\rangle$

$\equiv \langle 10, 0.7 \times 0.9; 100, 0.3 \times 0.7, 200, 0.7 \times 0.1, 1000, 0.3 \times 0.3 \rangle$

$E(S_1) = 0.3 \times [0.7 \times 100 + 0.3 \times 1000] + 0.7 \times [0.9 \times 10 + 0.1 \times 200]$

$= 0.3 \times E(\langle\langle D_2 = d_2 \rangle\rangle) + 0.7 \times E(\langle\langle D_3 = d_3 \rangle\rangle)$

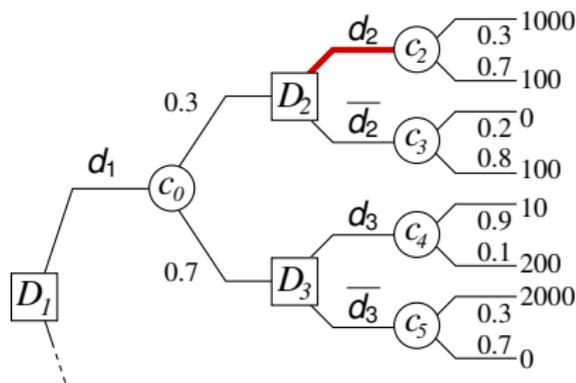
Stratégie  $S_2 = \langle\langle D_1 = d_1, D_2 = d_2, D_3 = \bar{d}_3 \rangle\rangle$

$\equiv \langle 0, 0.7 \times 0.7; 100, 0.3 \times 0.7, 1000, 0.3 \times 0.3, 2000, 0.7 \times 0.3 \rangle$

$E(S_2) = 0.3 \times [0.7 \times 100 + 0.3 \times 1000] + 0.7 \times [0.7 \times 0 + 0.3 \times 2000]$

$= 0.3 \times E(\langle\langle D_2 = d_2 \rangle\rangle) + 0.7 \times E(\langle\langle D_3 = \bar{d}_3 \rangle\rangle)$

# Calculs dans un arbre de décision (2/5)



Stratégie  $S_1 = \langle\langle D_1 = d_1, D_2 = d_2, D_3 = d_3 \rangle\rangle$

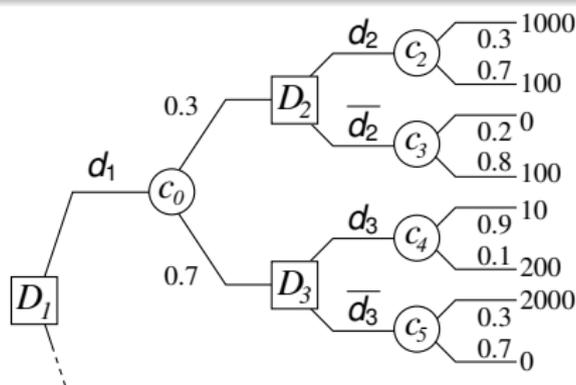
$$E(S_1) = 0.3 \times E(\langle\langle D_2 = d_2 \rangle\rangle) + 0.7 \times E(\langle\langle D_3 = d_3 \rangle\rangle)$$

Stratégie  $S_2 = \langle\langle D_1 = d_1, D_2 = \bar{d}_2, D_3 = \bar{d}_3 \rangle\rangle$

$$E(S_2) = 0.3 \times E(\langle\langle D_2 = \bar{d}_2 \rangle\rangle) + 0.7 \times E(\langle\langle D_3 = \bar{d}_3 \rangle\rangle)$$

$\Rightarrow$  calculer  $E(\langle\langle D_2 = d_2 \rangle\rangle)$  une seule fois, stocker le résultat en  $D_2$  et le réutiliser pour toute stratégie contenant  $D_2 = d_2$

# Calculs dans un arbre de décision (3/5)



**Problème :** Doit-on stocker en  $D_2$  les 2 espérances  $E(\langle\langle D_2 = d_2 \rangle\rangle)$  et  $E(\langle\langle D_2 = \bar{d}_2 \rangle\rangle)$  ?

Soit  $S_1 = \langle\langle D_1 = d_1, \dots, D_2 = d_2 \rangle\rangle$  et  $S_2 = \langle\langle D_1 = d_1, \dots, D_2 = \bar{d}_2 \rangle\rangle$   
deux stratégies ne différant que par la décision  $D_2$

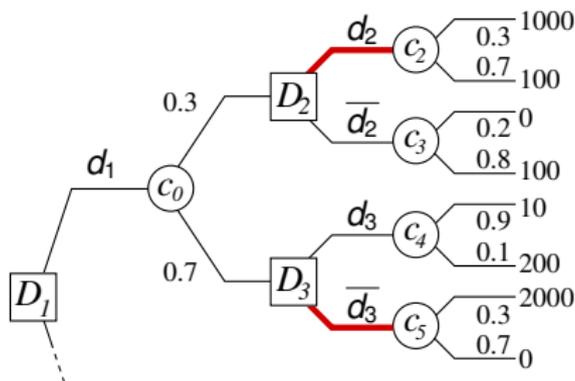
Alors  $E(S_1) - E(S_2) = E(\langle\langle D_2 = d_2 \rangle\rangle) - E(\langle\langle D_2 = \bar{d}_2 \rangle\rangle)$

$\implies$  Si  $E(\langle\langle D_2 = d_2 \rangle\rangle) \geq E(\langle\langle D_2 = \bar{d}_2 \rangle\rangle)$  alors  $E(S_1) \geq E(S_2)$

$\implies$  ne conserver que  $E(\langle\langle D_2 = d_2 \rangle\rangle)$  dans le nœud  $D_2$

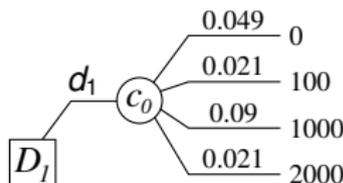
# Calculs dans un arbre de décision (4/5)

⇒ sur les nœuds de décision «terminaux», ne conserver que les meilleures décisions :



⇒ Si on choisit  $D_1 = d_1$ , la sous-stratégie optimale est forcément :  $D_2 = d_2, D_3 = d_3$  qui correspond à la loterie :

$\langle 0, 0.49; 100, 0.21; 1000, 0.09; 2000, 0.21 \rangle$

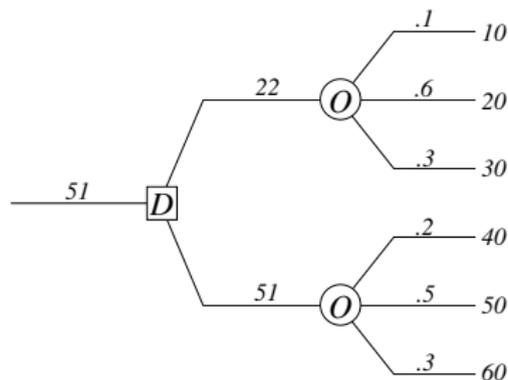


puis réitérer le process...

# Calculs dans un arbre de décision (5/5)

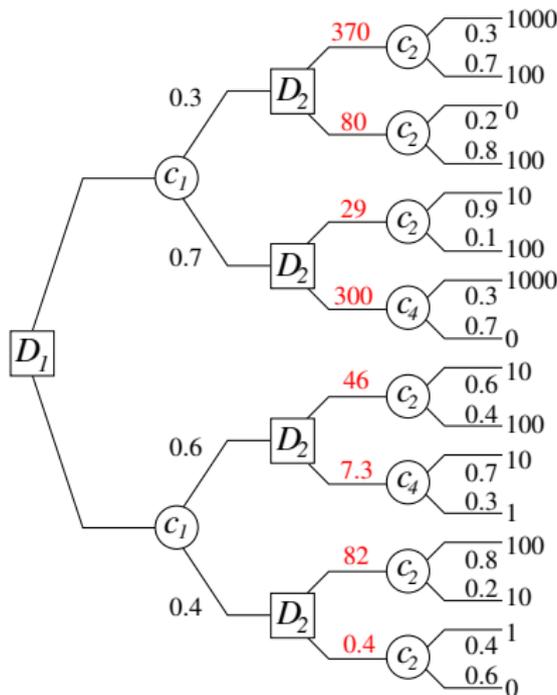
## Règle de calcul dans l'arbre de décision

- 1 si le nœud est un nœud de chance, on calcule une espérance
- 2 si le nœud est un nœud de décision, on conserve le max



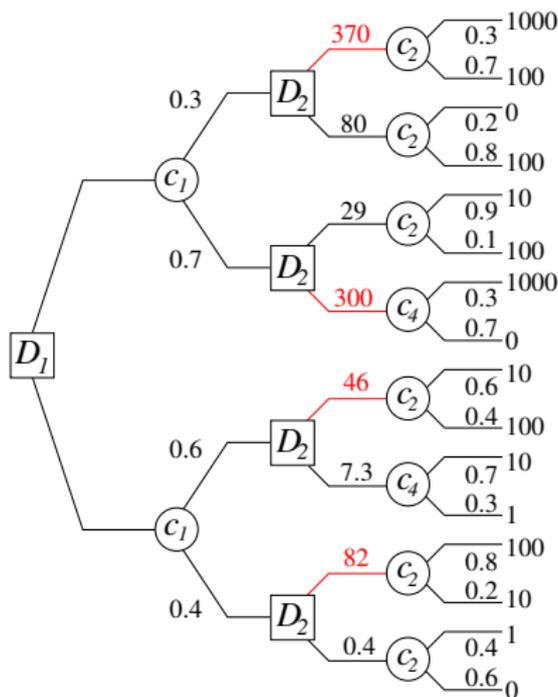
Méthode de calcul = inférence arrière

# Exemple d'inférence dans un arbre de décision (1/4)



$$\text{calcul : } EU(C_2) = \sum_{C_2} P(C_2|D_1, C_1, D_2)u(D_1, C_1, D_2, C_2)$$

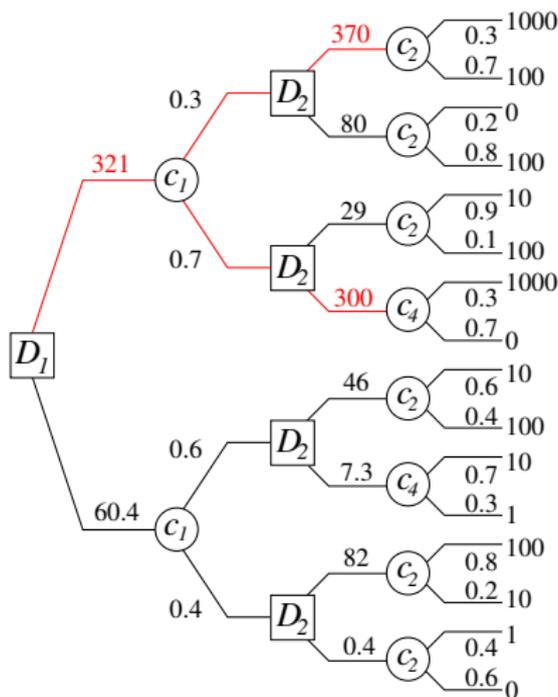
# Exemple d'inférence dans un arbre de décision (2/4)



$$\text{calcul : } EU(D_2) = \max_{D_2} EU(C_2)$$



# Exemple d'inférence dans un arbre de décision (4/4)

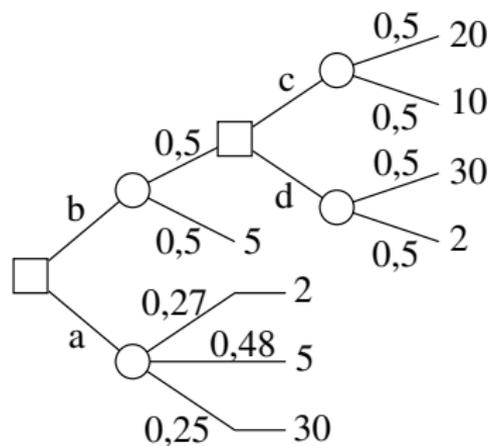


$$\text{calcul : } EU(D_1) = \max_{D_1} EU(C_1)$$

## ④ Décisions séquentielles avec RDU

# RDU : Problème de cohérence dynamique

Supposons que  $\varphi(x) = e^{-\sqrt{-\ln(x)}}$



$$RDU(a) = 2 + (5 - 2)\varphi(0,73) + (30 - 5)\varphi(0,25) = 11,41$$

$$RDU(bc) = 5 + (10 - 5)\varphi(0,5) + (20 - 10)\varphi(0,25) = 10,26$$

$$RDU(bd) = 2 + (5 - 2)\varphi(0,75) + (30 - 5)\varphi(0,25) = 11,46$$

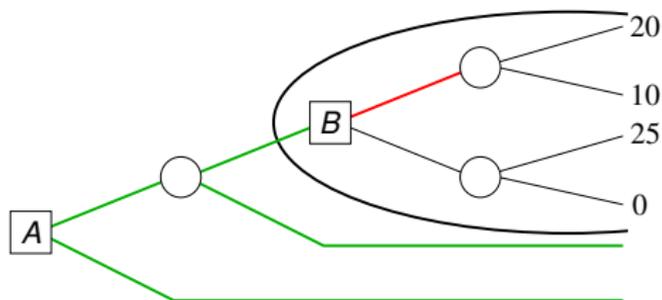
$$RDU(c) = 10 + (20 - 10)\varphi(0,5) = 14,35$$

$$RDU(d) = 2 + (30 - 2)\varphi(0,5) = 14,18$$

Problème : peut-on échapper au problème de cohérence dynamique ?

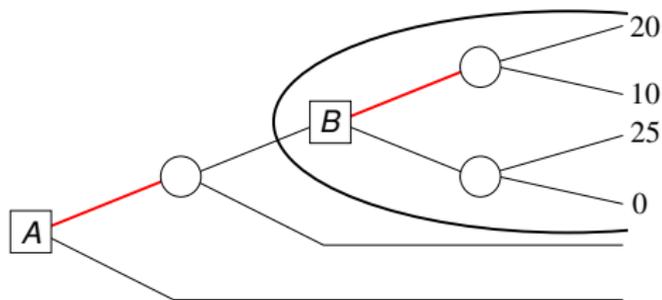
## Définition

Les préférences dans le sous-arbre de racine  $B$  ne dépendent pas du reste de l'arbre



## Définition

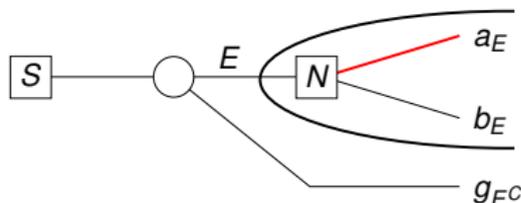
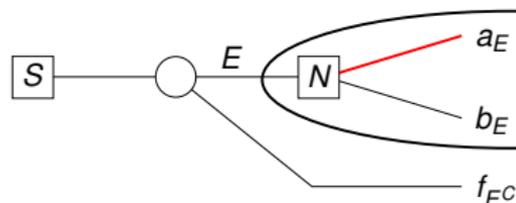
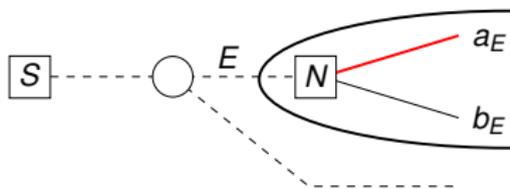
La stratégie préférée en  $A$  génère une sous-stratégie de racine  $B$  qui est la stratégie préférée en  $B$



# Conséquentialisme + cohérence dynamique

**conséquentialisme :**

pref ne dépend pas du  
sous-arbre en pointillé



**cohérence dynamique :** en  $S$ , la stratégie rouge en  $N$  est  
préférée dans les deux arbres

Donc  $(a_E, f_{EC}) \succ (b_E, f_{EC}) \iff (a_E, g_{EC}) \succ (b_E, g_{EC})$

conséquentialisme + cohérence dynamique  
 $\implies$  sure thing principle

## Références

- 1 **A. P. Dempster (1967)**  
“Upper and Lower Probabilities Induced by a Multivalued Mapping”, *Annals of Mathematical Statistics*, Vol 38, pp. 325-339
- 2 **G. Shafer (1976)**  
“Mathematical Theory of Evidence”, Princeton University Press
- 3 **J.-Y. Jaffray (1989)**  
“Linear Utility Theory For Belief Functions”, *Operations Research Letters*, Vol 8, pp. 107–112
- 4 **L. Hurwicz (1951)**  
“Optimality Criteria for Decision Making Under Ignorance”, Cowles Commission discussion paper, *Statistics*, N° 370