

# *Imaging et Sleeping Beauty*

Mikaël Cozic

DEC (ENS Ulm)

PhilEAs - 23/III/2007



## 1. Warm-Up: le changement de croyances

## dynamique des croyances

- ▶ *Sleeping Beauty* (SB) est un puzzle probabiliste qui déchaîne les passions en épistémologie formelle depuis qu'il a été introduit, en 2000, par A. Elga
- ▶ mon analyse de SB se focalise sur la *dynamique des croyances* dans le scénario

## dynamique des croyances

- ▶ *Sleeping Beauty* (SB) est un puzzle probabiliste qui déchaîne les passions en épistémologie formelle depuis qu'il a été introduit, en 2000, par A. Elga
- ▶ mon analyse de SB se focalise sur la *dynamique des croyances* dans le scénario
- ▶ 2 modèles principaux de changement de croyance:
  - (1) la **conditionalisation** pour les croyances partielles (*partial beliefs*)
  - (2) la **révision des croyances** pour les croyances “pleines” (*full beliefs*)

## les croyances partielles

- ▶ croyances partielles : Paul croit que  $A$  au degré  $r$
- ▶ le modèle orthodoxe de représentation des croyances partielles est le modèle *probabiliste*

## les croyances partielles

- ▶ croyances partielles : Paul croit que  $A$  au degré  $r$
- ▶ le modèle orthodoxe de représentation des croyances partielles est le modèle *probabiliste*
- ▶ les degrés de croyance obéissent aux axiomes des probabilités:
  1. si  $A$  est nécessairement vraie, alors  $P(A) = 1$   
(normalisation)
  2. pour toute  $A$ ,  $0 \leq P(A) \leq 1$  (non-négativité)
  3. si  $A$  et  $B$  sont incompatibles, alors  $P(A \vee B) = P(A) + P(B)$   
(additivité)

## exemple

- ▶ un panier peut contenir deux fruits, un ananas  $A$  ou une banane  $B$
- ▶  $W = \{AB, A\neg B, \neg AB, \neg A\neg B\}$ :

$AB$	$A\neg B$
$\neg AB$	$\neg A\neg B$

- ▶ Paul est certain qu'il y a au moins une fruit dans le panier et accorde la même probabilité aux 3 autres possibilités:

$1/3$	$1/3$
$1/3$	$0$

- ▶ de manière générale, (si  $W$  est fini)  $P(A) = \sum_{w \in A} P(w)$
- ▶ le **support** de  $P$  est l'ensemble des mondes qui ont un poids strictement positif:  $Supp(P) = \{w \in W : P(w) > 0\}$

## conditionalisation

- ▶ supposons que Paul apprenne une certaine information / (“il n’y a pas de banane dans le panier”); comment Paul doit-il faire évoluer ses croyances partielles ?



## conditionalisation

- ▶ supposons que Paul apprenne une certaine information  $I$  (“il n’y a pas de banane dans le panier”); comment Paul doit-il faire évoluer ses croyances partielles ?
- ▶ **conditionalisation:**  
$$P(A|I) = P(A \cap I) / P(I) = \sum_{w \in A \cap I} P(w) / \sum_{w' \in I} P(w')$$
- ▶  $P(A)$  était le poids des mondes qui sont des  $A$  relativement à l'ensemble  $W$  des mondes possibles;  $P(A|I)$  est le poids des mondes qui sont des  $A$  relativement aux mondes qui sont des  $I$

## exemple

- ▶ croyances initiales  $P$  de Paul :

$1/3$	$1/3$
$1/3$	$0$

- ▶ Paul apprend qu'il n'y a pas de banane i.e. apprend  $I = \{A \neg B, \neg A \neg B\}$  :

$0$	$1$
$0$	$0$

- ▶ Rem: *minimalité de la cond.* : soit  $w$  et  $w'$  des états compatibles avec l'information  $I$ . Alors les rapports entre  $P(w)$  et  $P(w')$  sont conservés par conditionalisation :  
$$P(w)/P(w') = P(w|I)/P(w'|I)$$

## Dutch Books

- ▶ pourquoi les degrés de croyances de Paul devraient-ils obéir aux axiomes des probabilités? Pourquoi devraient-ils évoluer conformément à la conditionalisation?
- ▶ justification **pragmatique** :
  - *Dutch Book statique* (Ramsey, de Finetti) pour la conformité aux axiomes des probabilités: si les degrés de croyances de Paul n'obéissent pas aux proba. et s'ils sont reflétés dans les paris que Paul est prêt à prendre, alors un *bookie* astucieux peut faire acheter à Paul un ensemble de paris tel que quel que soit le monde actuel, Paul en sort perdant
  - *Dutch Book dynamique* (Lewis, Teller) pour la conformité à la conditionalisation

## les croyances pleines

- ▶ croyances pleines: Paul croit que  $A$ /ne croit pas que  $A$ /croit que  $A$  n'est pas le cas
- ▶ les croyances de Paul sont représentées par un ensemble de mondes possibles  $K$  - les mondes compatibles avec ses croyances
- ▶ Paul
  - croit que  $A$  ssi  $A$  est vrai dans tous les mondes compatibles avec ses croyances - i.e. ssi  $K \subseteq A$
  - ne croit pas que  $A$  ssi  $A$  est faux dans certains des mondes compatibles avec ses croyances - i.e. ssi  $K \not\subseteq A$
  - croit que non  $A$  ssi  $A$  est faux dans tous les mondes compatibles avec ses croyances - i.e. ssi  $K \subseteq \bar{A}$

## exemple

- ▶ Paul croit qu'il y a au moins un fruit dans le panier *i.e.*  
 $K = \{AB, A\neg B, \neg AB\}$ :

$$\frac{\frac{AB \quad A\neg B}{\neg AB}}{\quad}$$

## révision

- ▶ supposons que Paul apprenne une certaine information / (“il n’y a pas de banane dans le panier”); comment Paul doit-il faire évoluer ses croyances (pleines)?

## révision

- ▶ supposons que Paul apprenne une certaine information  $I$  (“il n’y a pas de banane dans le panier”); comment Paul doit-il faire évoluer ses croyances (pleines)?
- ▶ croyances initiales  $K$ :

$$\frac{AB \quad A\neg B}{\neg AB}$$

- ▶ croyances révisées  $K^r I$ :

$$\frac{}{A\neg B}$$

- ▶ dans ce cas, l’information est **compatible** avec les croyances initiales:  $K \cap I \neq \emptyset$  et alors  $K^r I = K \cap I$  (on conserve les mondes qui peuvent l’être et on ajoute aucun monde supplémentaire)

## révision (2)

- ▶ la révision des croyances couvre également les cas où l'information n'est pas compatible avec les croyances initiales
  - ▶ axiomes de révision (AGM):
    - (A1) Si  $K \neq \emptyset$  et  $A \neq \emptyset$ ,  $K^r A \neq \emptyset$  (Cohérence)
    - (A2)  $K^r A \subseteq A$  (Succès)
    - (A3) Si  $K \subseteq A$ , alors  $K^r A = K$  (Conservation)
    - (A4)  $(K^r A) \cap B \subseteq K^r(A \cap B)$  (Sub-expansion)
    - (A5) Si  $(K^r A) \cap B \neq \emptyset$ , alors  $K^r(A \cap B) \subseteq (K^r A) \cap B$  (Super-expansion)
- De (A1)-(A4), on dérive :
- (A4')  $K \cap A \subseteq K^r A$  (Inclusion)
- De (A1)-(A5), on dérive :
- (A5') Si  $K \cap A \neq \emptyset$ , alors  $K^r A \subseteq K \cap A$  (Préservation')



## révision (3)

- ▶ on peut rendre précise l'idée que les axiomes codifient une forme de changement minimal de croyances
- ▶ théorème de représentation: un opérateur de révision  $(.\prime.)$  satisfait (A1)-(A5) ssi il existe une famille  $\leq_K$  de relations de similitudes sur les mondes telle que  $K^r A = \{w' \in A: \forall w'' \in A, w' \leq_K w''\}$
- ▶ autrement dit, les mondes du nouvel ensemble de croyances sont les mondes qui satisfont l'information  $(A)$  et qui sont les plus proches du point de vue des croyances initiales

## pour résumer...

- ▶ la conditionalisation
  - (i) s'applique aux croyances partielles
  - (ii) ne couvre pas les informations incompatibles
  - (iii) est justifiée pragmatiquement
- ▶ la révision
  - (i)\* s'applique aux croyances pleines
  - (ii)\* couvre les informations incompatibles
  - (iii)\* n'est pas justifiée pragmatiquement

## 2. SB, Halfers & Thirders

## le scénario

- ▶ le dimanche soir ( $t_0$ ), SB est endormie
- ▶ le lundi matin ( $t_1$ ), SB est réveillée (l'expérimentateur ne lui dit pas quel jour on est)
- ▶ quelques minutes après ( $t_2$ ), on dit à SB que l'on est lundi
- ▶ la suite dépend du tirage d'une pièce non biaisée qui a lieu le dimanche soir. SB ne connaît pas l'issue du tirage.
- ▶ si *PILE*, SB est réveillée le mardi matin puis rendormie jusqu'à la fin de la semaine. Mais on a donné à SB une drogue qui fait qu'elle est incapable de savoir si l'on est lundi ou mardi.
- ▶ si *FACE*, SB est endormie après qu'on lui ait dit qu'on était lundi ( $t_2$ ) et jusqu'à la fin de la semaine

## 2 questions

- ▶ ce à quoi l'on s'intéresse, c'est la probabilité que SB devrait accorder au fait que le résultat du tirage de la pièce a été *FACE*
- ▶ 2 questions
  - Q1 quelle devrait être selon SB la probabilité de *FACE* à  $t_1$  ?
  - Q2 quelle devrait être selon SB la probabilité de *FACE* à  $t_2$  ?
- ▶ Notation:
  - $P_0$  = croyances de SB à  $t_0$  (dimanche soir)
  - $P_1$  = croyances de SB à  $t_1$  (à son réveil le lundi matin)
  - $P_2$  = croyances de SB à  $t_2$  (le lundi matin après qu'on lui ait dit qu'on était lundi)

## Halfers et Thirders

- ▶ Thèse des Thirders (Elga, 2000):  $P_1(FACE) = 1/3$
- ▶ Thèse des Halfers (Lewis, 2001):  $P_1(FACE) = 1/2$
- ▶ **Mais** les réponses à Q1 sont en réalité étroitement liées aux réponses à Q2 et les deux positions sont mieux décrites de la manière suivante :

	A. Elga	D. Lewis
Q1	1/3	1/2
Q2	1/2	2/3

## prémises communes

- ▶ espace d'états (interprétés comme "mondes centrés")

$W = \{FL, PL, PM\}$  où

- $FL$  la pièce tombe sur face et l'on est lundi
- $PL$  la pièce tombe sur pile et l'on est lundi
- $PM$  la pièce tombe sur pile et l'on est mardi

- ▶ *prémises communes* à Elga et Lewis:

$$(1) P_1(PL) = P_1(PM)$$

(principe d'indifférence ou de Laplace)

$$(2) P_2(FACE) = P_1(FACE|LUNDI) = P_1(FACE|FL \cup PL)$$

(changement de croyances par conditionalisation)

$$(5) P_0(FACE) = P_0(PILE) = 1/2$$

( $\approx$  Principe Principal de Lewis)

## l'argument d'Elga

- ▶ idée de base : la pièce pourrait être lancée le lundi soir. Dans ce cas, que serait  $P_2(FACE)$  ? La même que le dimanche soir soit

$$(E) P_2(FACE) = P_0(FACE) = 1/2$$

- ▶ De (E) et (1), (2), (5) il suit que  $P_1(FACE) = 1/3$  en "remontant" temporellement la conditionalisation :

$$P_2(FACE) = P_1(FACE|FL \cup PL) = 1/2.$$

- ▶ argument "Bottom-Up" qui répond à Q1 en répondant d'abord à Q2. L'argument use du Principe Principal pour établir  $P_2(FACE)$ .

Problème : s'agit-il d'un usage licite du Principe Principal ?



## l'argument de Lewis

- ▶ idée de base : quand SB se réveille le lundi matin ( $t_1$ ), elle n'acquiert aucune "évidence non centrée" pertinente pour l'issue *PILE/FACE*. Par conséquent ses croyances relativement à *PILE/FACE* doivent rester inchangées entre  $P_0$  et  $P_1$ :  
(L)  $P_1(\text{FACE}) = 1/2 = P_1(\text{PILE})$
- ▶ De (L) et (1), (2), (5), il suit que  
 $P_2(\text{FACE}) = 2/3$
- ▶ argument "Top-Down" qui répond à Q2 en répondant à Q1. L'argument use d'un principe d'inertie doxastique selon lequel ma croyance en  $P$  doit rester inchangée si je n'apprends rien de *pertinent* relativement à  $P$ .  
Problème : selon quel critère décide-t-on qu'une information est pertinente relativement à une proposition ?

## SB généralisée

- ▶ désormais, si *PILE*, alors SB est réveillée  $N$  jours (lundi inclus) tandis que si *FACE*, SB n'est réveillée que le lundi.

- ▶ argument d'Elga :

$$P_2(\text{FACE}) = 1/2, \text{ donc}$$

$$P_1(\text{FACE}) = 1/(N + 1)$$

$$\text{Pour } N = 100, P_1(\text{FACE}) \approx 0.0099$$

Quand SB se réveille, elle est presque sûre que la pièce est tombée sur *PILE*.

- ▶ argument de Lewis :

$$P_1(\text{FACE}) = 1/2, \text{ donc}$$

$$P_2(\text{FACE}) = N/(N + 1)$$

$$\text{Pour } N = 100, P_2(\text{FACE}) \approx 0.99$$

Quand SB apprend que l'on est lundi, elle est presque sûre que la pièce est tombée sur *FACE*.

## 3. Conditionalisation & imaging

## point de départ

- ▶ double constat :
  - 1) les intuitions de départ de Lewis et Elga sont toutes les deux attrayantes. Si on acceptait ces intuitions conjointement, on aboutirait à une position *double halfer* selon laquelle
$$P_1(FACE) = P_2(FACE) = 1/2$$
  - 2) étant donné les prémisses communes, ces deux intuitions sont incompatibles
- ▶ pourquoi ? Elga et Lewis assignent tous deux une probabilité de  $1/2$  à *FACE* : Lewis à  $t_1$ , Elga à  $t_2$ . Etant donné la prémisse commune de conditionalisation entre  $t_1$  et  $t_2$  (sur l'événement *LUNDI*), ces probabilités ne peuvent pas être identiques au même moment.

## conditionalisation (1)

- ▶ la situation serait différente si l'on supposait une autre règle de changement de croyances que la conditionalisation. Mais a-t-on des raisons de remettre en cause la conditionalisation ?

## conditionalisation (1)

- ▶ la situation serait différente si l'on supposait une autre règle de changement de croyances que la conditionalisation. Mais a-t-on des raisons de remettre en cause la conditionalisation ?
- ▶ la proposition qu'apprend SB à  $t_2$  porte sur sa *localisation temporelle*
- ▶ elle est par conséquent *sensible au contexte* (sa valeur de vérité dépend du contexte temporel)
- ▶ les propositions sensibles au contexte temporel sont problématiques pour la conditionalisation
- ▶ il y a deux propriétés de la conditionalisation qui posent problème :
  - (i) la concentration
  - (ii) la partialité

## conditionalisation (2)

- (i) concentration : la conditionalisation ne peut que *concentrer* les croyances initiales : la probabilité d'une proposition  $P$  peut baisser, mais la probabilité d'un monde compatible avec l'information ne peut qu'augmenter.

Si l'information  $I$  est compatible avec les croyances initiales  $P$  ( $I \cap \text{Supp}(P) \neq \emptyset$ ), alors  $\text{Supp}(P(.|I)) \subseteq \text{Supp}(P)$

► Cas particuliers :

- si une proposition  $A$  est certaine et si elle est compatible avec l'information, alors elle reste certaine (*preservation*, Gardenförs (1988))
- si un monde a probabilité nulle, il ne peut plus jamais acquérir une probabilité négative.

## conditionalisation (3)

- ▶ **SB** : (a) la probabilité de  $FL$  ne peut qu'augmenter quand SB apprend que l'on est lundi (b) si, à  $t_0$ , SB croit que l'on est dimanche, elle ne peut pas, à  $t_1$  croire que l'on est lundi ou mardi
- (ii) partialité : la conditionalisation n'est pas définie quand l'information est incompatible avec les croyances initiales ( $I \cap \text{Supp}(P) = \emptyset$ )

**SB** : la conditionalisation n'indique pas comment passer de  $P_0$  à  $P_1$  : à  $t_0$  le support de  $P_0$  est, disons,  $\{FD, PD\}$ , mais à  $t_1$  SB "apprend"  $I_1 = \{FL, PL, PM\}$ .



## conditionalisation et SB

- ▶ ces propriétés suggèrent que quand des propositions sensibles au contexte sont en jeu, la conditionalisation n'est pas un guide fiable pour le changement de croyances
- ▶ d'où l'idée de chercher une alternative à la façon dont les Halfers et les Thirders usent de la conditionalisation.
- ▶ 2 grande options :
  - 1) conserver la règle de conditionalisation mais amender le cadre sous-jacent (Bostrom, Halpern)
  - 2) conserver le cadre sous-jacent mais se fier à un autre règle que la conditionalisation
- ▶ quelle autre règle ?

## imaging

- ▶ Lewis (1976) introduit la règle d'*imaging*. Soit une proposition  $A \subseteq W$ . On se donne comme notion primitive la notion de *monde le plus proche de  $w$  où  $A$  est vraie*, notée  $w_A$  - (i) Si  $w \in A$ ,  $w_A = w$ . (ii)  $w_A \in A$ .  
cf. la sémantique de Stalnaker pour les conditionnels
- ▶ Supposons que l'on apprenne que  $A$ . Alors la règle d'*imaging* dit que la probabilité d'un monde  $w$  est transférée au monde  $w_A$ .

$$P^{Im(A)}(w) = \sum_{\{w' \in W : w = w'_A\}} P(w')$$

- ▶ Lewis: "no gratuitous movement of probability from worlds to dissimilar worlds"

## exemple

- ▶ le panier de fruits:

$AB$	$A\neg B$
$\neg AB$	$\neg A\neg B$

- ▶ probabilité initiale  $P$ :

$1/3$	$1/3$
$1/3$	$0$

- ▶ changement par imaging sur  $I = \{A\neg B, \neg A\neg B\}$  avec  $AB_I = A\neg B$  et  $\neg AB_I = \neg A\neg B$ :

$0$	$2/3$
$0$	$1/3$

## l'imaging est-il sérieux ?

- ▶ l'*imaging* n'est en général pas considérée comme une règle sérieuse de changement de croyances.
- ▶ Lewis (1976) l'introduite parce qu'il correspond à la probabilité des conditionnels de Stalnaker.
- ▶ Gardenförs (1988) l'écarte car elle viole la propriété de *preservation*.
- ▶ **mais** il existe une tentative récente de justification de (l'usage de) l'*imaging* : Walliser & Zwirn (2002).
- ▶ **idée fondamentale** : la conditionalisation est appropriée dans un certain type de contextes (révision), l'imaging dans un autre (mise à jour)

## 4. Révision & mise à jour

## contextes de changement de croyance

- ▶ dans les théories formelles du changement de croyances "pleines", on distingue deux contextes de changement de croyance :
  - 1) les contextes de *révision* : l'agent apprend une nouvelle information à propos d'un environnement supposé stable
  - 2) les contextes de *mise à jour* : l'agent apprend un changement dans l'environnement

## contextes de changement de croyance

- ▶ dans les théories formelles du changement de croyances "pleines", on distingue deux contextes de changement de croyance :
  - 1) les contextes de *révision* : l'agent apprend une nouvelle information à propos d'un environnement supposé stable
  - 2) les contextes de *mise à jour* : l'agent apprend un changement dans l'environnement
- ▶ Katzuno & Mendelzon (1992) ont fait valoir que le second type de contexte requiert des principes différents du premier.

## exemple

- ▶ ensemble initial de croyances  $K = \{AB, A\neg B, \neg AB\}$ :

$$\frac{AB \quad A\neg B}{\neg AB}$$

- ▶ information  $I = \{A\neg B, \neg A\neg B\}$

### revising

"il n'y a pas de banane"

$$\frac{A\neg B}{\quad}$$

### updating

"il n'y a plus de banane (s'il y en avait une)"

$$\frac{A\neg B}{\neg A\neg B}$$



## axiomes de rationalité

- ▶ (A1)-(A4) sont communs (donc aussi l'axiome d'inclusion :  
(A4')  $K \cap A \subseteq K^r A$
- ▶ l'axiome propre à la révision est  
(A5) Si  $(K^r A) \cap B \neq \emptyset$ , alors  $K^r(A \cap B) \subseteq (K^r A) \cap B$   
(Super-expansion)
- ▶ les axiomes propres à la mise à jour sont :  
(A6) Si  $\exists w_0 \in W$  t.q.  $K = \{w_0\}$  et si  $(K^m A) \cap B \neq \emptyset$ , alors  
 $K^m(A \cap B) \subseteq (K^m A) \cap B$  (Super-expansion pointée)
- (A7)  $(K \cup K')^m A = (K^m A) \cup (K'^m A)$  (Distributivité gauche)  
(A7) "gives each of the possible worlds equal consideration" (KM 1992)

## relations de similitude

- (i) Révision : famille  $\leq_K$  de similitudes sur les mondes.

$$K^r A = \{w' \in A \text{ t.q. } \forall w'' \in A, w' \leq_K w''\}$$

$\Rightarrow$  changement minimal *global*

- (ii) Mise à jour : famille  $\leq_w$  de similitudes sur les mondes.

$$K^m A = \{w' \in A \text{ t.q. } \exists w \in K \text{ et } \forall w'' \in A : w' \leq_w w''\}$$

$\Rightarrow$  changement minimal *local*

## 5. Mise à jour, imaging & SB

## justification de l'imaging

- ▶ Katzuno & Mendelzon (1992) suggèrent qu'il y a une correspondance entre la conditionalisation et la révision d'une part, et l'imaging et la mise à jour d'autre part
- ▶ Walliser & Zwirn (2002) ont clarifié cette suggestion en montrant que :
  - (i) les règles de changement de croyances de type conditionalisation sont dérivables d'une transcription probabiliste des axiomes AGM pour la révision des croyances
  - (ii) les règles de changement de croyances de type imaging sont dérivables d'une transcription probabiliste des axiomes KM pour la mise à jour des croyances

## *imaging* et SB

- ▶ **idée**: quand SB apprend que l'on est lundi ( $t_2$ ), elle a une information sur une caractéristique de sa situation qui a *changé* depuis ses croyances initiales
- ▶ il semble par conséquent plus approprié de modéliser le changement de croyances de SB par la règle d'*imaging*
- ▶ pour appliquer l'*imaging*, il faut supposer une relation de similitude entre mondes. Quelle relation de similitude ?
- ▶ hypothèse : *PL* est le monde-*LUNDI* le plus proche de *PM*.

## l'argument pour les *Double-Halfers*

- ▶ point de départ de Lewis :

$$P_1(\text{FACE}) = P_1(\text{FL}) = 1/2 \text{ et } P_1(\text{PL}) = P_1(\text{PM}) = 1/4$$

Par *imaging*, on obtient

$$P_2(\text{PL}) = P_1(\text{PL}) + P_1(\text{PM}) = 1/2$$

$$P_2(\text{FACE}) = P_2(\text{FL}) = P_1(\text{FL}) = 1/2$$

- ▶ point de départ d'Elga :

$$P_2(\text{FACE}) = P_2(\text{FL}) = 1/2 = P_2(\text{PILE}) = P_2(\text{PL})$$

En prenant l'*imaging* à rebours on obtient

$$P_1(\text{FACE}) = P_2(\text{FACE}) = 1/2$$

## 6. Discussion

## l'argument général

- ▶ résumé de l'argument général:
  - (P1) dans un contexte (probabiliste) de mise à jour, il faut changer ses croyances par *imaging* et non par conditionalisation
  - (P2) quand SB apprend que l'on est lundi, il s'agit d'un contexte de mise à jour
  - (C) SB doit changer ses croyances par *imaging* quand elle apprend que l'on est lundi



## l'objection principale

- ▶ quand SB se réveille, à  $t_1$ , et qu'elle prend conscience que l'on est lundi ou mardi, il s'agit d'un authentique contexte de mise à jour (on est plus le même jour qu'en  $t_0$ )  
**Mais** quand SB apprend que l'on est lundi à  $t_2$ , ce qu'elle apprend *n'est pas* un changement qui aurait eu lieu entre  $t_1$  et  $t_2$
- ▶ par conséquent, même si l'*imaging* était approprié dans les contextes de mise à jour, elle ne s'appliquerait pas en  $t_2$  car il ne s'agit pas d'un contexte de mise à jour
- ▶ **problème général**: quand un agent apprend successivement deux informations  $I_1, I_2$  à deux instants  $t_1 < t_2$  à propos d'un changement qui a eu lieu entre  $t_0$  et  $t_1$ , est-ce que  $I_2$  doit être traitée par conditionalisation ou par *imaging* ?

## exemple: Ananas, Banane & Cerise

- ▶ espace d'états

$ABC$	$AB\neg C$	$A\neg BC$	$A\neg B\neg C$
$\neg ABC$	$\neg AB\neg C$	$\neg A\neg BC$	$\neg A\neg B\neg C$

- ▶ probabilité initiale à  $t_0$ :

0	1/4	1/4	1/4
1/4	0	0	0

- ▶ information  $I_1 = \{A\neg BC, A\neg B\neg C, \neg A\neg BC, \neg A\neg B\neg C\}$

- ▶ *imaging* sur  $I_1$  :  $P_1 = P_0^{Im(I_1)}$

0	0	1/4	1/2
0	0	1/4	0

## exemple, cont.

- ▶ information  $I_2 = \{A \neg B \neg C, \neg A \neg B \neg C\}$  (“il n’y a plus de banane (s’il y en avait une) et il n’y a plus de cerise (s’il y en avait une)”) à  $t_2$
- ▶ *imaging* sur  $I_2$  :

0	0	0	3/4
0	0	0	1/4

Rem:  $P_0^{Im(I_2)} = P_1^{Im(I_2)}$

- ▶ conditionalisation sur  $I_2$ :

0	0	0	1
0	0	0	0

Rem:  $P_0^{Cond(I_2)} = P_1^{Cond(I_2)}$

## analyse

- ▶ dans l'exemple comme dans SB, le changement a lieu entre  $t_0$  et  $t_1$
- ▶ dans les deux cas, la seconde information  $I_2$  porte sur un changement qui a eu lieu entre  $t_0$  et  $t_1$  et *raffine* l'information  $I_1$  (i.e.  $I_2 \subset I_1$ )
- ▶ si l'on est convaincu par la distinction entre révision et mise par jour par un exemple comme Ananas & Banane, on doit préférer  $P_1^{Im(I_2)}$  à  $P_1^{Cond(I_2)}$  dans Ananas, Banane & Cerise

## conclusion

- ▶ *l'imaging* ouvre la voie à une authentique réconciliation entre les intuitions fondamentales des Halfers et des Thirders
- ▶ la distinction entre révision et mise à jour n'est pas assez nette pour étayer un argument définitif dans un scénario comme SB
- ▶ les règles de changement de croyances dépendent de paramètres temporels qui ne sont pas explicités
- ▶ un modèle satisfaisant de changement de croyances devrait expliciter de tels paramètres

## Dutch Book dynamique

- ▶  $P_0$  distribution équiprobable sur  
 $Supp(P_0) = \{AB, \neg AB, A\neg B\}$ .  $I = \{\neg A\neg B, A\neg B\}$ .
- ▶  $P^{Im(I)}(AB) = P^{Im(I)}(\neg AB) = 0$   $P^{Im(I)}(A\neg B) = 2/3$ . On suppose que  $P_1 = P^{Im(I)}$ .
- ▶ notons que  $P^{Im(I)}(A) = 2/3$  alors que  $P(A|I) = 1$
- ▶ Trois paris :
  - (a) [1 si  $A \wedge \neg B$ , 0 sinon]
  - (b) [ $x = P_0(A|I) = 1$  si  $B$ , 0 sinon]
  - (c) [ $y = P_0(A|I) - P_1(A) = 1/3$  si  $\neg B$ , 0 sinon]L'agent les achète à  $1/3 + 2/3 + 1/9 = 11/9$ .

## Dutch Book dynamique

- ▶ Cas 1 : il est faux que  $I = \neg B$ , alors l'agent a une perte sèche de  $y.P_0(I) = 1/3.1/3 = 1/9$
- ▶ Cas 2 : on informe l'agent que  $I = \neg B$ , alors dans le *Dutch Book* diachronique, le *bookie* lui rachète le pari (d) [1 si A, 0 sinon] pour  $P_1(A)$ . L'agent essuie aussi une perte sèche de  $y.P_0(I) = 1/3.1/3 = 1/9$ .
- ▶ Problème : le rachat du pari.  
Pari (a) porte sur : il est vrai à  $t_0$  que A et que  $\neg B$   
Pari (d) porte sur : il est vrai à  $t_1$  que A