Réseaux causaux possibilistes pour le traitement des interventions

Salem BENFERHAT
CRIL, Lens
benferhat@cril.univ-artois.fr

Les modèles graphiques probabilistes

Outils importants pour la représentation
et l'analyse des informations incertaines
dans les systèmes de base de connaissances

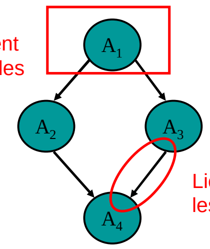
2

Réseaux Bayésiens

- Composante « symbolique »
  - Structure graphique
    - Graphe acyclique orienté (DAG)
    - Influence et Indépendance
- Composante quantitative
  - distributions de probabilités
    - Incertitude
    - Indépendance

3

Réseaux bayésiens (RB)

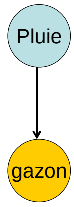


Graphe acyclique orienté
+ Distributions de probabilité conditionnelles locales

4

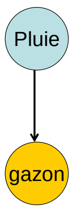
Observations vs. Interventions

Dépendance



- Si on observe que le gazon est mouillé
- P(gazon= « mouillé »)=1
- est-ce que P(Pluie) ↑ ou ↓?
- Oui

Causalité



- On mouille le gazon
- P(gazon= « mouillé »)=1
- est-ce que P(Pluie) ↑ ou ↓?
- Non

Intervention: interne ou externe?

Une intervention est d’origine externe

Réseau bayésien causal

Réseau bayésien causal

- Réseau bayésien causal (RBC) est :
  - un réseau bayésien, mais :
    - Un arc $X \rightarrow Y$ existe ssi X est une cause directe de Y
    - Arc $\leftrightarrow$ Lien de causalité

➡ P(X|Y) signifie que non seulement
il existe *une relation entre X et Y* mais que Y est une cause X

Réseau bayésien causal

- Un réseau bayésien causal prend un sens uniquement si on effectue des interventions.

Intervention

- « Le gazon est mouillé » => observation
=> conditionnement probabiliste: $P(\omega) \rightarrow P(\omega \mid A=a)$

- « On mouille le gazon » => intervention ou action
Comment modéliser les interventions dans un RB?

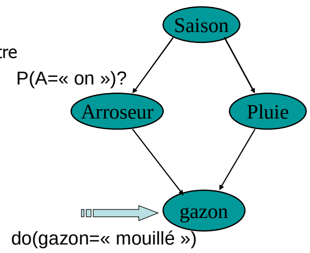
- Pearl (1992):
  - Une action qui force A à prendre la valeur a est notée do(A=a) ou do(a).

$$P(\omega) \rightarrow P(\omega \mid do(A=a))$$

Intervention

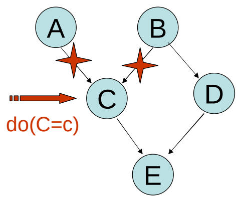
- Après intervention sur Arroseur (A),

⇒ Les croyances sur les causes directes de A ne doivent pas être changées



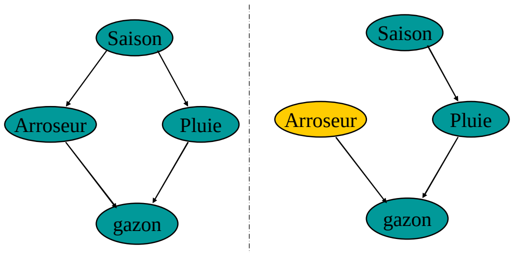
Représentation de l 'intervention (1): « Mutilation du graphe »

- Supprimer les liens vers le nœud concerné
- Toutes les causes autres que celle de l 'intervention sont exclues
- Conditionnement sur le graphe modifié (mutilé).



- Les croyances sur A et B ne changent pas

Représentation de l 'intervention (1): « Mutilation du graphe »



Observation : **A=« On »**
Conditionnement classique

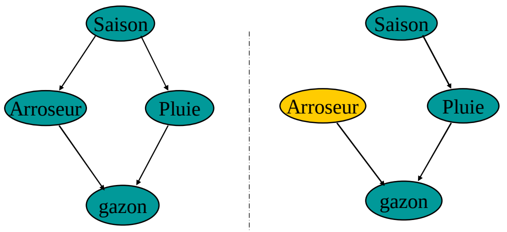
Action : do(**A=« On »**)
- supprimer les arcs entrant vers A
- Conditionnement dans le nouveau graphe

Calcul de l 'effet d'une action

- Cas : pas de variables non-observées (non-mesurées)
- Action simple *do(A=a$_i$')*
- *Effet de l'intervention sur la jointe est:*

- $P(a_1,\ldots,a_i,\ldots, a_n \mid do(a_i')) = \begin{cases} \prod_{j \neq i} P(a_j \mid u_{Aj}) & \text{si } a_i = a_i', \\ 0 & \text{sinon} \end{cases}$

Calcul de l 'effet d 'une action



P(SAPH)=

P(S) P(A|S) P(P|S) P(H|AP)

P(SAPH)=

P(S) P(P|S) P(H|AP)

Algo de propagation sur les interventions =

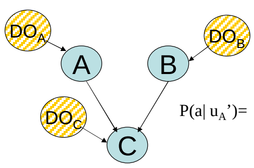
Algo de propagation sur les observations

17

Autres représentations équivalents des interventions

Représentation de l'intervention (2): « Augmentation du graphe »

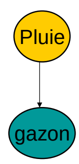
- Do(A=a')
- Pour toute variable A où il peut y avoir interventions, ajouter un nœud parent DO_A
- DO_A est une variable qui prend valeur dans {do(A=a'),rien} pour tout a' instance de A.
- « rien » signifie « pas d 'actions ».
- $U_A \leftarrow \{U_A \cup DO_A\}$



$$P(a|\ u_A') = \begin{cases} P(a|\ u_A) & \text{si } DO_A=\text{rien} \\ 1 & \text{si } DO_A=do(a') \text{ et } a'=a \\ 0 & \text{si } DO_A=do(a') \text{ et } a'\neq a \end{cases}$$

19

Exemple = graphe augmenté



Pluie	
oui	0,8
non	0,2

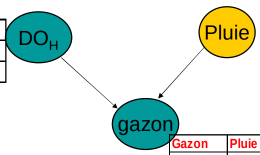
Gazon	Pluie	Herbe|Pluie
mouillé	oui	0,9
mouillé	non	0,3
sec	oui	0,1
sec	non	0,7

20

Exemple = graphe augmenté

?

rien	
do(mouillée)	
do(sèche)	



Pluie	
oui	0,8
non	0,2

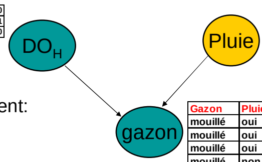
Gazon	Pluie	DO_H	Herbe|Pluie,DO_H
mouillé	oui	rien	0,9
mouillé	oui	do(mouillé)	1
mouillé	oui	do(sec)	0
mouillé	non	rien	0,3
mouillé	non	do(mouillé)	1
mouillé	non	do(sec)	0
sec	oui	rien	0,1
sec	oui	do(mouillé)	1
sec	oui	do(sec)	0
sec	non	rien	0,7
sec	non	do(mouillé)	1
sec	non	do(sec)	0

21

Exemple = graphe augmenté

Si il y a intervention do(gazon=mouillé)

rien	0
do(mouillée)	1
do(sèche)	0



Pluie	
oui	0,8
non	0,2

Gazon	Pluie	DO_H	Herbe|Pluie,DO_H
mouillé	oui	rien	0,9
mouillé	oui	do(mouillé)	1
mouillé	oui	do(sec)	0
mouillé	non	rien	0,3
mouillé	non	do(mouillé)	1
mouillé	non	do(sec)	0
sec	oui	rien	0,1
sec	oui	do(mouillé)	0
sec	oui	do(sec)	1
sec	non	rien	0,7
sec	non	do(mouillé)	0
sec	non	do(sec)	1

on obtient:

Gazon	Pluie	DO_H	P(Herbe,Pluie,DO_H)
mouillé	oui	rien	0
mouillé	oui	do(mouillé)	0,8
mouillé	oui	do(sec)	0
mouillé	non	rien	0
mouillé	non	do(mouillé)	0,2
mouillé	non	do(sec)	0
sec	oui	rien	0
sec	oui	do(mouillé)	0
sec	oui	do(sec)	0
sec	non	rien	0
sec	non	do(mouillé)	0
sec	non	do(sec)	0

22

Graphe augmenté

- Objectif :

Graphe augmenté

Arbre de jonction

Inférence sur interventions

23

Graphe augmenté: propagation

- Problème : quelle distribution a priori attribuer aux nœuds DO_{Ai} (à l'état initial)?
- Ce que l'on veut exprimer est:
  - Par défaut il n'y a pas d'intervention
  - Qui peut-être remise en cause s'il y a intervention

24

Graphe augmenté: propagation

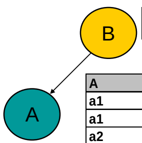
- Par défaut il n'y a pas d'intervention
  - si on attribue 1 à $P(DO_{Ai}$=« rien »)
  - ⇒On ne peut plus après avoir DO_{Ai}=«a_i'»
  - ⇒(on est en présence d'événement impossible)

Graphe augmenté: propagation

- Donc il faut faire une initialisation avec des probabilité positive sur chacune des instance de DO_{Ai}
- Par exemple, la solution de distribution uniforme sur la variable DO_{Ai}
  - $P(DO_{Ai}$=«rien»)=1/k+1
  - Pour tout ai, $P(DO_{Ai}=a_i)$=1/k+1

Exemple : probabilité

Observation B=b1



b1	1
b2	0

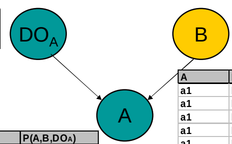
A	B	A|B
a1	b1	0,6
a1	b2	0,3
a2	b1	0,4
a2	b2	0,7

A	B	P(AB)
a1	b1	0,6
a1	b2	0
a2	b1	0,4
a2	b2	0

- P(A=a1)=0.6
- P(A=a2)=0.4

Exemple : probabilité

rien	0,333
do(a1)	0,333
do(a2)	0,333



b1	1
b2	0

A	B	DOA	A|B,DOA
a1	b1	rien	0,6
a1	b1	do(a1)	1
a1	b1	do(a2)	0
a1	b2	rien	0,3
a1	b2	do(a1)	1
a1	b2	do(a2)	0
a2	b1	rien	0,4
a2	b1	do(a1)	0
a2	b1	do(a2)	1
a2	b2	rien	0,7
a2	b2	do(a1)	0
a2	b2	do(a2)	1

A	B	DOA	$P(A,B,DO_A)$
a1	b1	rien	0,2
a1	b1	do(a1)	0,333
a1	b1	do(a2)	0
a1	b2	rien	0
a1	b2	do(a1)	0
a1	b2	do(a2)	0
a2	b1	rien	0,133
a2	b1	do(a1)	0
a2	b1	do(a2)	0,333
a2	b2	rien	0
a2	b2	do(a1)	0
a2	b2	do(a2)	0

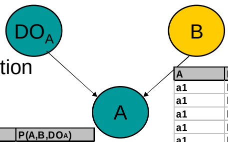
- P(A=a1)= 0.5333
- P(A=a2)= 0.4666

=> P(A) =/= $P_{GA}(A)$

Exemple : probabilité

rien	0,333
do(a1)	0,333
do(a2)	0,333

Pas d 'intervention



b1	1
b2	0

A	B	DOA	A|B,DOA
a1	b1	rien	0,6
a1	b1	do(a1)	1
a1	b1	do(a2)	0
a1	b2	rien	0,3
a1	b2	do(a1)	1
a1	b2	do(a2)	0
a2	b1	rien	0,4
a2	b1	do(a1)	0
a2	b1	do(a2)	1
a2	b2	rien	0,7
a2	b2	do(a1)	0
a2	b2	do(a2)	1

A	B	DOA	P(A,B,DOA)
a1	b1	rien	0,2
a1	b1	do(a1)	0,333
a1	b1	do(a2)	0
a1	b2	rien	0
a1	b2	do(a1)	0
a1	b2	do(a2)	0
a2	b1	rien	0,133
a2	b1	do(a1)	0
a2	b1	do(a2)	0,333
a2	b2	rien	0
a2	b2	do(a1)	0
a2	b2	do(a2)	0

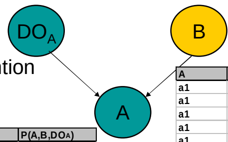
- P(A=a1)= 0.5333
- P(A=a2)= 0.4666

=> P(A) =/= P_{GA}(A)

Exemple : probabilité

rien	1
do(a1)	0
do(a2)	0

Pas d 'intervention



b1	1
b2	0

A	B	DOA	A|B,DOA
a1	b1	rien	0,6
a1	b1	do(a1)	1
a1	b1	do(a2)	0
a1	b2	rien	0,3
a1	b2	do(a1)	1
a1	b2	do(a2)	0
a2	b1	rien	0,4
a2	b1	do(a1)	0
a2	b1	do(a2)	1
a2	b2	rien	0,7
a2	b2	do(a1)	0
a2	b2	do(a2)	1

A	B	DOA	P(A,B,DOA)
a1	b1	rien	0,6
a1	b1	do(a1)	0
a1	b1	do(a2)	0
a1	b2	rien	0
a1	b2	do(a1)	0
a1	b2	do(a2)	0
a2	b1	rien	0,4
a2	b1	do(a1)	0
a2	b1	do(a2)	0
a2	b2	rien	0
a2	b2	do(a1)	0
a2	b2	do(a2)	0

- P(A=a1)= 0.6
- P(A=a2)= 0.4

=> P(A) = P_{GA}(A)

Graphe augmenté: propagation

- On peut faire une étape de transformation vers arbre de jonction, mais avant de propager, il faut connaître toutes les interventions.
- Solutions possibles dans d'autres cadres de l'incertain

Réseaux causaux possibilistes

Motivations

- ✓ L'opérateur de base « do » a été introduit dans le cadre des OCF (Fonctions Conditionnelles Ordinales)
- ✓ La notion de causalité a été motivée en utilisant uniquement la structure de graphe
  Ex: Une intervention sur une variable ne doit pas changer nos croyances sur ses causes directes.
- ✓ Il est clair que l'intervention et l'observation ne doivent pas être confondues.

=> donc il est important d'avoir un cadre qui permet, en théorie de possibilité, de modéliser les interventions.

33

Réseau causal possibiliste

- Réseau causal possibiliste (RCP) est:
  - un réseau possibiliste, mais :
    - Un arc X→Y existe ssi X est une cause directe de Y
    - Arc ↔ Lien de causalité

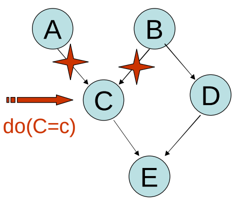
π(X|Y) signifie que non seulement il existe *une relation entre X et Y* mais que Y est une cause de X

- Un réseau bayésien causal prend un sens uniquement si on effectue des interventions.

34

Représentation de l 'intervention (1): « Mutilation du graphe »

- Supprimer les liens vers le nœud concerné
- Toutes les causes autres que celle de l 'intervention sont exclues
- Conditionnement sur le graphe modifié (mutilé) selon le type de réseaux (basés sur le min ou sur le produit)

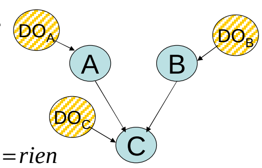


- Les croyances sur A et B ne changent pas:
  π(A)=π'(A)=π(A|do(C=c))
  π(B)=π'(B)=π(B|do(C=c))

35

Représentation de l'intervention (2): « Augmentation du graphe »

- Do(A=a')
- Pour toute variable A où il peut y avoir interventions, ajouter un nœud parent DO_A
- DO_A est une variable qui prend valeur dans {do(A=a'),rien} pour tout a' instance de A.
- « rien » signifie « pas d 'actions ».
- $U_A \leftarrow \{U_A \cup DO_A\}$

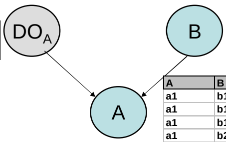


$$\pi(a \mid u_A') = \begin{cases} \pi(a \mid u_A) & \text{si } do_A = rien \\ 1 & \text{si } do_A = do(a') \text{ et } a' = a \\ 0 & \text{si } do_A = do(a') \text{ et } a' \neq a \end{cases}$$

36

Graphe augmenté: avantages

rien	?
do(a1)	?
do(a2)	?



b1	1
b2	0,4

A	B	DOA	A|B,DOA
a1	b1	rien	1
a1	b1	do(a1)	1
a1	b1	do(a2)	0
a1	b2	rien	0,6
a1	b2	do(a1)	1
a1	b2	do(a2)	0
a2	b1	rien	0,5
a2	b1	do(a1)	0
a2	b1	do(a2)	1
a2	b2	rien	1
a2	b2	do(a1)	0
a2	b2	do(a2)	1

- <u>Problème:</u> quelle distribution a priori attribuer aux nœuds DO_A(à l'état initial)?
- Ce que l'on veut exprimer est:
  - pas d'intervention par défaut
  - possibilité de la remise en cause s'il y a intervention

37

Graphe augmenté : propagation

- En probabilité:

Pas d'intervention par défaut => $\forall$Ai, $P(DO_{Ai}$=« rien »)=1

$\Rightarrow$ On ne peut plus après avoir DO_{Ai}=«a_i'»
(évènement impossible)

- Par exemple, la solution de distribution uniforme sur la variable DO_{Ai}
  - $P(DO_{Ai}$=«rien»)=1/$|D_{Ai}|$+1
  - $\forall a_i$, $P(DO_{Ai}=a_i)=1/|D_{Ai}|+1$

38

Graphe augmenté possibiliste: propagation

- Solution: cadre possibiliste
- $\forall$Ai, distributions de possibilité locales au niveau de DO_{Ai}:
  - $\pi(DO_{Ai}$=rien) $\leftarrow$ 1
  - $\forall a_i$' instance de Ai, $\pi(DO_{Ai}$=do(ai')) $\leftarrow \varepsilon$
    où α est le plus bas degré de possibilité attribué (toutes variables considérées)

$$\varepsilon \leq \min_{ai}(\pi(a_i|u_{Ai}))$$

- $\forall$Ai, DO_{Ai}=rien est considérée *par défaut* comme étant la situation la plus normale
- $\forall a_i$', DO_{Ai}=do(ai') est considérée comme étant l'évènement *le moins normal et le moins préféré* dans le graphe *évitant ainsi d'influencer nos connaissances initiales* sur le reste des variables

39

Graphe augmenté possibiliste: propagation

- Distributions par défaut équivalentes sur les variables de base

$$\forall \omega\ \Pi_{GA}(\omega)= \pi(\omega)$$

- En absence d'intervention, la distribution de possibilité associée au graphe augmenté est égale à la distribution de possibilité associée au grahe initial

40

Arbres de jonction augmentés

- Les nœuds ajoutés DO_{Ai} au graphe initial G sont aussi ajoutés à l'arbre de jonction AJ

$\Rightarrow$ Construction de AJ à partir du graphe augmenté GA : problématique !

Nbre de nœuds = Nbre nœuds dans G+Nbre nœuds ajoutés

- Solution: Ajouter les nœuds DO_{Ai} lors de l'étape d'initialisation de AJ

41

Arbres de jonction augmentés

- Initialisation:
  - Pour chaque clique C_i dans AJ:

    $\forall\omega,\ \pi_{Ci}(\omega)\leftarrow 1$
  - Pour chaque A_i dans G, choisir une clique C_i contenant $A_i \cup U_{Ai}$
    - Si A_i est un nœud qui peut être concerné par une intervention:
      - Ajouter le nœud DO_{Ai} à C_i
      - $\pi_{Ci} \leftarrow \pi_{Ci} \,.\, \pi(a_i|u_{Ai}, DO_{Ai}) \,.\, \pi(DO_{Ai})$
    - Sinon
      - $\pi_{Ci} \leftarrow \pi_{Ci} \,.\, \pi(a_i|u_{Ai})$

42

Arbres de jonction augmentés

- La jointe résultante dans AJ est équivalente à celle du graphe augmenté GA:

$$\forall\omega, do_i,\ \pi_{AJ}(\omega, do_i)=\pi_{GA}(\omega, do_i)$$

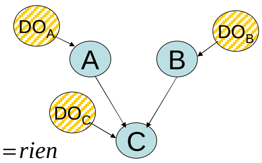
où do_i est une instance quelconque des nœuds DO_{Ai}

43

Représentation de l'intervention (2): « Augmentation du graphe »

Représentation de l'intervention (2): « Augmentation du graphe »

- Pour toute variable A où il peut y avoir interventions, ajouter un nœud parent DO_A
- DO_A est une variable qui prend valeur dans {do(A=a'),rien} pour tout a' instance de A.
- « rien » signifie « pas d 'interventions ».
- $U_A \leftarrow \{U_A \cup DO_A\}$



$$\pi(a \mid u_A^{'}) = \begin{cases} \pi(a \mid u_A) & \text{si } do_A = rien \\ 1 & \text{si } do_A = do(a^{'}) \text{ et } a^{'} = a \\ 0 & \text{si } do_A = do(a^{'}) \text{ et } a^{'} \neq a \end{cases}$$

45

Inférence dans des graphes augmentés possibilistes : Problème de constructions d'arbres de jonction pour chaque intervention

46

Inférence dans les réseaux possibilistes

- Observation C=c

- Effet de cette observation sur V\{C} ?
  - π(A|C=c), π(B|C=c),……

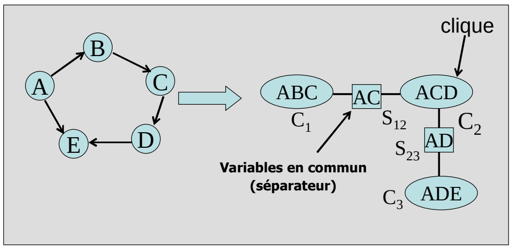
  => Inférence

47

Propagation sans intervention dans les RP

- RB simplement connectés : polynomial
- RB à connexions multiples : NP-complet

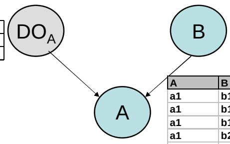
⇒ Transformation graphique : graphe initial en un arbre de jonction (moralisation - triangulation)



48

Graphe augmenté: retour en probabilités

rien	?
do(a1)	?
do(a2)	?



b1	1
b2	0,4

A	B	DOA	A|B,DOA
a1	b1	rien	1
a1	b1	do(a1)	1
a1	b1	do(a2)	0
a1	b2	rien	0,6
a1	b2	do(a1)	1
a1	b2	do(a2)	0
a2	b1	rien	0,5
a2	b1	do(a1)	0
a2	b1	do(a2)	1
a2	b2	rien	1
a2	b2	do(a1)	0
a2	b2	do(a2)	1

- <u>Problème:</u> (ignorance sur la présence ou non d'interventions) quelle distribution a priori attribuer aux nœuds DO_A(à l'état initial)?

- Ce que l'on veut exprimer est:
  - pas d'intervention par défaut
  - possibilité de la remise en cause s'il y a intervention

49

Graphe augmenté : propagation

- En probabilité:

Pas d'intervention par défaut => $\forall$Ai, P(DO_{Ai}=« rien »)=1

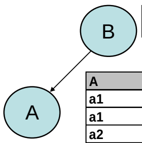
$\Rightarrow$ On ne peut plus après avoir DO_{Ai}=«a_i'»
(évènement impossible)

- La solution de distribution uniforme sur la variable DO_{Ai} n'est pas satisfaisante non plus.
  - P(DO_{Ai}=«rien»)=1/|D_{Ai}|+1
  - $\forall a_i$, P(DO_{Ai}=a_i)=1/|D_{Ai}|+1

50

Graphe augmenté (probabilités)

Observation B=b1



b1	1
b2	0

A	B	A|B
a1	b1	0,6
a1	b2	0,3
a2	b1	0,4
a2	b2	0,7

<u>Distribution jointe P</u>

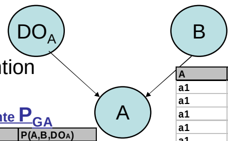
A	B	P(AB)
a1	b1	0,6
a1	b2	0
a2	b1	0,4
a2	b2	0

- P(A=a1)=0.6
- P(A=a2)=0.4

51

Graphe augmenté (probabilités)

rien	0,333
do(a1)	0,333
do(a2)	0,333



b1	1
b2	0

Pas d 'intervention

<u>Distribution jointe P_{GA}</u>

A	B	DOA	P(A,B,DOA)
a1	b1	rien	0,2
a1	b1	do(a1)	0,333
a1	b1	do(a2)	0
a1	b2	rien	0
a1	b2	do(a1)	0
a1	b2	do(a2)	0
a2	b1	rien	0,133
a2	b1	do(a1)	0
a2	b1	do(a2)	0,333
a2	b2	rien	0
a2	b2	do(a1)	0
a2	b2	do(a2)	0

A	B	DOA	A|B,DOA
a1	b1	rien	0,6
a1	b1	do(a1)	1
a1	b1	do(a2)	0
a1	b2	rien	0,3
a1	b2	do(a1)	1
a1	b2	do(a2)	0
a2	b1	rien	0,4
a2	b1	do(a1)	0
a2	b1	do(a2)	1
a2	b2	rien	0,7
a2	b2	do(a1)	0
a2	b2	do(a2)	1

- P(A=a1)= 0.5333
- P(A=a2)= 0.4666

=> **$P(A) \neq P_{GA}(A)$**

52

Graphe augmenté possibiliste: propagation

- 2ème solution:
  - $\pi(DO_{Ai}=rien) \leftarrow 1$
  - $\forall a_i'$ instance de Ai, $\pi(DO_{Ai}=do(ai')) \leftarrow \alpha$
    où α est le plus bas degré de possibilité attribué (toutes variables considérées)

- $\forall$Ai, DO_{Ai}=rien est considérée *par défaut* comme étant la situation la plus normale
- $\forall a_i'$, DO_{Ai}=do(ai') est considérée comme étant l'évènement *le moins normal et le moins préféré* dans le graphe *évitant ainsi d'influencer nos connaissances initiales* sur le reste des variables

FIN