

# Apprentissage par renforcement de PDM factorisés avec effets corrélés

JFPDA 2010

Boris LESNER et Bruno ZANUTTINI

GREYC, Université de Caen Basse-Normandie

1 juin 2010

- 1 Introduction
- 2 Effets connus à l'avance
- 3 Effets non observables
- 4 Conclusion

On s'intéresse aux domaines :

- Représentables de manière compacte avec des Opérateurs STRIPS Probabilistes.
  - **Généralisation** de l'expérience.
  - Gestion du compromis Exploration/Exploitation.
- Avec peu d'hypothèses sur la structure.
  - La structure est **inconnue** *à priori*.
  - On connaît seulement sa **taille**.

## Apprentissage de la politique - *model free*

- Apprentissage très indirect.
- Utilisation directe.

# Les différentes approches

## Apprentissage de la politique - *model free*

- Apprentissage très indirect.
- Utilisation directe.

## Apprentissage de la fonction de valeur - *Q-learning*

- Apprentissage semi-direct (Q-valeurs).
- Utilisation moins directe (action qui maximise).

# Les différentes approches

## Apprentissage de la politique - *model free*

- Apprentissage très indirect.
- Utilisation directe.

## Apprentissage de la fonction de valeur - *Q-learning*

- Apprentissage semi-direct (Q-valeurs).
- Utilisation moins directe (action qui maximise).

## Apprentissage du modèle - *model based*

- Apprentissage direct (transitions et récompenses).
- Utilisation indirecte (résoudre un PDM).

# Exemple introductif

## Jouer au loto.

- 3 variables d'état
  - être **chanceux** :  $c$ , ou non :  $\bar{c}$
  - être **riche** :  $r$ , ou non :  $\bar{r}$
  - être un **vendredi 13** :  $v$ , ou non :  $\bar{v}$
- Les chanceux qui jouent au loto ont une plus grande probabilité de gagner que les autres.
- Le vendredi 13 ne change rien.

| Condition $x$ | Effet $e$ | Proba $T(e s \models x)$ |
|---------------|-----------|--------------------------|
| $c$           | $r$       | 0.2                      |
|               | $\bar{r}$ | 0.8                      |
| $\bar{c}$     | $r$       | 0.1                      |
|               | $\bar{r}$ | 0.9                      |

# Exemple introductif

## Jouer au loto.

- 3 variables d'état
  - être **chanceux** :  $c$ , ou non :  $\bar{c}$
  - être **riche** :  $r$ , ou non :  $\bar{r}$
  - être un **vendredi 13** :  $v$ , ou non :  $\bar{v}$
- Les chanceux qui jouent au loto ont une plus grande probabilité de gagner que les autres.
- Le vendredi 13 ne change rien.

| Condition $x$ | Effet $e$ | Proba $T(e s \models x)$ |
|---------------|-----------|--------------------------|
| $c$           | $r$       | 0.2                      |
|               | $\bar{r}$ | 0.8                      |
| $\bar{c}$     | $r$       | 0.1                      |
|               | $\bar{r}$ | 0.9                      |

Seul le fait d'être chanceux influe sur la probabilité de gagner.

# La difficulté d'apprendre ces actions

- On ne connaît pas la structure des conditions :
  - ni leur nombre,
  - ni quelles variables interviennent (mais on sait combien)
- Dans un PDM, l'agent n'observe que des états : quel effet à provoqué chaque transition ?

## Reconnaitre les effets du Loto

- Une personne **non riche** ( $\bar{r}$ ) joue et gagne, dans l'état suivant on observe bien l'effet  $r$ .
- Une personne **déjà riche** ( $r$ ) joue et gagne, dans l'état suivant on observe aucun changement, pourtant l'effet  $r$  s'est produit.

# Exemple Introductif (suite)

Apprendre à jouer au loto.

- On sait que **une seule** variable apparaît dans les conditions de l'action.
  - mais on ne connaît pas laquelle.

Jouer au loto dépend soit :

- de  $c$ , avec les conditions  $c$  et  $\bar{c}$  ;
- soit de  $r$ , avec  $r$  et  $\bar{r}$ .
- ou encore de  $v$ , avec  $v$  et  $\bar{v}$ .

On maintient donc les trois hypothèses simultanément.

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |  |
|-----------|--|
| $c$       |  |
| $\bar{c}$ |  |

Hypothèse 2 :  $r$

|           |  |
|-----------|--|
| $r$       |  |
| $\bar{r}$ |  |

Hypothèse 3 :  $v$

|           |  |
|-----------|--|
| $v$       |  |
| $\bar{v}$ |  |

On observe

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |  |
|-----------|--|
| $c$       |  |
| $\bar{c}$ |  |

Hypothèse 2 :  $r$

|           |  |
|-----------|--|
| $r$       |  |
| $\bar{r}$ |  |

Hypothèse 3 :  $v$

|           |  |
|-----------|--|
| $v$       |  |
| $\bar{v}$ |  |

On observe

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |          |
|-----------|----------|
| $c$       | $\top$ 1 |
| $\bar{c}$ |          |

Hypothèse 2 :  $r$

|           |          |
|-----------|----------|
| $r$       |          |
| $\bar{r}$ | $\top$ 1 |

Hypothèse 3 :  $v$

|           |          |
|-----------|----------|
| $v$       | $\top$ 1 |
| $\bar{v}$ |          |

On observe  $(c\bar{r}v, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $\top$ | 1 |
| $\bar{c}$ | $\top$ | 1 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       |        |   |
| $\bar{r}$ | $\top$ | 2 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $\top$ | 1 |
| $\bar{v}$ | $\top$ | 1 |

On observe  $(\bar{c}\bar{r}\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |          |
|-----------|----------|
| $c$       | $\top$ 2 |
| $\bar{c}$ | $\top$ 1 |

Hypothèse 2 :  $r$

|           |          |
|-----------|----------|
| $r$       | $\top$ 1 |
| $\bar{r}$ | $\top$ 2 |

Hypothèse 3 :  $v$

|           |          |
|-----------|----------|
| $v$       | $\top$ 2 |
| $\bar{v}$ | $\top$ 1 |

On observe  $(crv, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $\top$ | 2 |
| $\bar{c}$ | $\top$ | 2 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $\top$ | 2 |
| $\bar{r}$ | $\top$ | 2 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $\top$ | 2 |
| $\bar{v}$ | $\top$ | 2 |

On observe  $(\bar{c}r\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |          |
|-----------|----------|
| $c$       | $\top$ 3 |
| $\bar{c}$ | $\top$ 2 |

Hypothèse 2 :  $r$

|           |          |
|-----------|----------|
| $r$       | $\top$ 2 |
| $\bar{r}$ | $\top$ 3 |

Hypothèse 3 :  $v$

|           |          |
|-----------|----------|
| $v$       | $\top$ 3 |
| $\bar{v}$ | $\top$ 2 |

On observe  $(c\bar{r}v, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |              |
|-----------|--------------|
| $c$       | $r$ 1<br>T 3 |
| $\bar{c}$ | T 2          |

Hypothèse 2 :  $r$

|           |              |
|-----------|--------------|
| $r$       | T 2          |
| $\bar{r}$ | $r$ 1<br>T 3 |

Hypothèse 3 :  $v$

|           |              |
|-----------|--------------|
| $v$       | T 3          |
| $\bar{v}$ | $r$ 1<br>T 2 |

On observe  $(c\bar{r}\bar{v}, r)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $r$    | 1 |
|           | $\top$ | 3 |
| $\bar{c}$ |        |   |
|           | $\top$ | 3 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $\top$ | 2 |
|           | $r$    | 1 |
| $\bar{r}$ |        |   |
|           | $\top$ | 4 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $\top$ | 3 |
|           | $r$    | 1 |
| $\bar{v}$ |        |   |
|           | $\top$ | 3 |

On observe  $(\bar{c}\bar{r}\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |               |        |
|-----------|---------------|--------|
| $c$       | $r$<br>$\top$ | 1<br>4 |
| $\bar{c}$ | $\top$        | 3      |

Hypothèse 2 :  $r$

|           |               |        |
|-----------|---------------|--------|
| $r$       | $\top$        | 3      |
| $\bar{r}$ | $r$<br>$\top$ | 1<br>4 |

Hypothèse 3 :  $v$

|           |               |        |
|-----------|---------------|--------|
| $v$       | $\top$        | 3      |
| $\bar{v}$ | $r$<br>$\top$ | 1<br>4 |

On observe  $(c\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $r$    | 1 |
|           | $\top$ | 4 |
| $\bar{c}$ | $r$    | 1 |
|           | $\top$ | 3 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 1 |
|           | $\top$ | 3 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 4 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $\top$ | 3 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 4 |

On observe  $(\bar{c}r\bar{v}, r)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

## Hypothèse 1 : $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $r$    | 2 |
|           | $\top$ | 4 |
| $\bar{c}$ | $r$    | 1 |
|           | $\top$ | 3 |

## Hypothèse 2 : $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 2 |
|           | $\top$ | 3 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 4 |

## Hypothèse 3 : $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $r$    | 1 |
|           | $\top$ | 3 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 4 |

On observe  $(crv, r)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $r$    | 2 |
|           | $\top$ | 4 |
| $\bar{c}$ | $r$    | 1 |
|           | $\top$ | 4 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 2 |
|           | $\top$ | 3 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 5 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $r$    | 1 |
|           | $\top$ | 3 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 5 |

On observe  $(\bar{c}\bar{r}\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $r$    | 2 |
|           | $\top$ | 4 |
| $\bar{c}$ | $r$    | 1 |
|           | $\top$ | 5 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 2 |
|           | $\top$ | 4 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 5 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $r$    | 1 |
|           | $\top$ | 3 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 6 |

On observe  $(\bar{c}r\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $r$    | 2 |
|           | $\top$ | 4 |
| $\bar{c}$ | $r$    | 1 |
|           | $\top$ | 6 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 2 |
|           | $\top$ | 4 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 6 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $r$    | 1 |
|           | $\top$ | 3 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 7 |

On observe  $(\bar{c}\bar{r}\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $r$    | 2 |
|           | $\top$ | 5 |
| $\bar{c}$ | $r$    | 1 |
|           | $\top$ | 6 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 2 |
|           | $\top$ | 4 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 7 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $r$    | 1 |
|           | $\top$ | 4 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 7 |

On observe  $(c\bar{r}v, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $r$    | 2 |
|           | $\top$ | 5 |
| $\bar{c}$ | $r$    | 1 |
|           | $\top$ | 7 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 2 |
|           | $\top$ | 4 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 8 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $r$    | 1 |
|           | $\top$ | 5 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 7 |

On observe  $(\bar{c}\bar{r}v, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $r$    | 2 |
|           | $\top$ | 6 |
| $\bar{c}$ | $r$    | 1 |
|           | $\top$ | 7 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 2 |
|           | $\top$ | 5 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 8 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $r$    | 1 |
|           | $\top$ | 6 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 7 |

On observe  $(crv, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $r$    | 2 |
|           | $\top$ | 6 |
| $\bar{c}$ | $r$    | 1 |
|           | $\top$ | 8 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 2 |
|           | $\top$ | 6 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 8 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $r$    | 1 |
|           | $\top$ | 7 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 7 |

On observe  $(\bar{c}rv, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |   |
|-----------|--------|---|
| $c$       | $r$    | 2 |
|           | $\top$ | 6 |
| $\bar{c}$ | $r$    | 1 |
|           | $\top$ | 9 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 2 |
|           | $\top$ | 7 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 8 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $r$    | 1 |
|           | $\top$ | 8 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 7 |

On observe  $(\bar{c}rv, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 2  |
|           | $\top$ | 6  |
| $\bar{c}$ | $r$    | 1  |
|           | $\top$ | 10 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 2 |
|           | $\top$ | 7 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 9 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $r$    | 1 |
|           | $\top$ | 9 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 7 |

On observe  $(\bar{c}\bar{r}v, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 2  |
|           | $\top$ | 6  |
| $\bar{c}$ | $r$    | 1  |
|           | $\top$ | 11 |

Hypothèse 2 :  $r$

|           |        |   |
|-----------|--------|---|
| $r$       | $r$    | 2 |
|           | $\top$ | 8 |
| $\bar{r}$ | $r$    | 1 |
|           | $\top$ | 9 |

Hypothèse 3 :  $v$

|           |        |   |
|-----------|--------|---|
| $v$       | $r$    | 1 |
|           | $\top$ | 9 |
| $\bar{v}$ | $r$    | 2 |
|           | $\top$ | 8 |

On observe  $(\bar{c}r\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 2  |
|           | $\top$ | 6  |
| $\bar{c}$ | $r$    | 1  |
|           | $\top$ | 12 |

Hypothèse 2 :  $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 2  |
|           | $\top$ | 8  |
| $\bar{r}$ | $r$    | 1  |
|           | $\top$ | 10 |

Hypothèse 3 :  $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 1  |
|           | $\top$ | 10 |
| $\bar{v}$ | $r$    | 2  |
|           | $\top$ | 8  |

On observe  $(\bar{c}\bar{r}v, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 2  |
|           | $\top$ | 6  |
| $\bar{c}$ | $r$    | 1  |
|           | $\top$ | 13 |

Hypothèse 2 :  $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 2  |
|           | $\top$ | 9  |
| $\bar{r}$ | $r$    | 1  |
|           | $\top$ | 10 |

Hypothèse 3 :  $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 1  |
|           | $\top$ | 11 |
| $\bar{v}$ | $r$    | 2  |
|           | $\top$ | 8  |

On observe  $(\bar{c}rv, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |     |    |
|-----------|-----|----|
| $c$       | $r$ | 3  |
|           | T   | 6  |
| $\bar{c}$ | $r$ | 1  |
|           | T   | 13 |

Hypothèse 2 :  $r$

|           |     |    |
|-----------|-----|----|
| $r$       | $r$ | 2  |
|           | T   | 9  |
| $\bar{r}$ | $r$ | 2  |
|           | T   | 10 |

Hypothèse 3 :  $v$

|           |     |    |
|-----------|-----|----|
| $v$       | $r$ | 2  |
|           | T   | 11 |
| $\bar{v}$ | $r$ | 2  |
|           | T   | 8  |

On observe  $(c\bar{r}v, r)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 3  |
|           | $\top$ | 6  |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 13 |

Hypothèse 2 :  $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 2  |
|           | $\top$ | 9  |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 10 |

Hypothèse 3 :  $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 11 |
| $\bar{v}$ | $r$    | 2  |
|           | $\top$ | 8  |

On observe  $(\bar{c}\bar{r}v, r)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |     |    |
|-----------|-----|----|
| $c$       | $r$ | 4  |
|           | T   | 6  |
| $\bar{c}$ | $r$ | 2  |
|           | T   | 13 |

Hypothèse 2 :  $r$

|           |     |    |
|-----------|-----|----|
| $r$       | $r$ | 3  |
|           | T   | 9  |
| $\bar{r}$ | $r$ | 3  |
|           | T   | 10 |

Hypothèse 3 :  $v$

|           |     |    |
|-----------|-----|----|
| $v$       | $r$ | 3  |
|           | T   | 11 |
| $\bar{v}$ | $r$ | 3  |
|           | T   | 8  |

On observe  $(c\bar{v}, r)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 7  |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 13 |

Hypothèse 2 :  $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 9  |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 11 |

Hypothèse 3 :  $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 12 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 8  |

On observe  $(c\bar{r}v, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 7  |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 14 |

Hypothèse 2 :  $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 10 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 11 |

Hypothèse 3 :  $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 13 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 8  |

On observe  $(\bar{c}rv, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

## Hypothèse 1 : $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 7  |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 15 |

## Hypothèse 2 : $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 10 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 12 |

## Hypothèse 3 : $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 13 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 9  |

On observe  $(\bar{c}\bar{r}\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

## Hypothèse 1 : $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 8  |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 15 |

## Hypothèse 2 : $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 10 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 13 |

## Hypothèse 3 : $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 13 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 10 |

On observe  $(c\bar{r}\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 9  |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 15 |

Hypothèse 2 :  $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 11 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 13 |

Hypothèse 3 :  $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 14 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 10 |

On observe  $(crv, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 10 |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 15 |

Hypothèse 2 :  $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 11 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 14 |

Hypothèse 3 :  $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 14 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 11 |

On observe  $(c\bar{r}\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 11 |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 15 |

Hypothèse 2 :  $r$

|           |        |    |
|-----------|--------|----|
|           | $r$    | 3  |
| $r$       | $\top$ | 11 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 15 |

Hypothèse 3 :  $v$

|           |        |    |
|-----------|--------|----|
|           | $r$    | 3  |
| $v$       | $\top$ | 14 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 12 |

On observe  $(c\bar{r}\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

## Hypothèse 1 : $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 12 |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 15 |

## Hypothèse 2 : $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 12 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 15 |

## Hypothèse 3 : $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 15 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 12 |

On observe  $(crv, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 13 |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 15 |

Hypothèse 2 :  $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 13 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 15 |

Hypothèse 3 :  $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 15 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 13 |

On observe  $(c\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 13 |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 16 |

Hypothèse 2 :  $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 13 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 16 |

Hypothèse 3 :  $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 16 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 13 |

On observe  $(\bar{c}\bar{r}v, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

## Hypothèse 1 : $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 14 |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 16 |

## Hypothèse 2 : $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 14 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 16 |

## Hypothèse 3 : $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 16 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 14 |

On observe  $(c\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

## Hypothèse 1 : $c$

|           |                    |
|-----------|--------------------|
| $c$       | $r$ 4<br>$\top$ 15 |
| $\bar{c}$ | $r$ 2<br>$\top$ 16 |

## Hypothèse 2 : $r$

|           |                    |
|-----------|--------------------|
| $r$       | $r$ 3<br>$\top$ 15 |
| $\bar{r}$ | $r$ 3<br>$\top$ 16 |

## Hypothèse 3 : $v$

|           |                    |
|-----------|--------------------|
| $v$       | $r$ 3<br>$\top$ 16 |
| $\bar{v}$ | $r$ 3<br>$\top$ 15 |

On observe  $(c\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

## Hypothèse 1 : $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 16 |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 16 |

## Hypothèse 2 : $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 15 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 17 |

## Hypothèse 3 : $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 16 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 16 |

On observe  $(c\bar{r}\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

Hypothèse 1 :  $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 16 |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 17 |

Hypothèse 2 :  $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 16 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 17 |

Hypothèse 3 :  $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 16 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 17 |

On observe  $(\bar{c}r\bar{v}, \top)$

# Mise à jour des hypothèses

Au cours de l'apprentissage on observe des transitions de la forme :  $(s, e)$ , où  $e$  est l'effet qui est appliqué dans l'état  $s$ .

- On va donc classer ces exemples dans les différentes hypothèses sur le modèle.

## Hypothèse 1 : $c$

|           |        |    |
|-----------|--------|----|
| $c$       | $r$    | 4  |
|           | $\top$ | 16 |
| $\bar{c}$ | $r$    | 2  |
|           | $\top$ | 18 |

## Hypothèse 2 : $r$

|           |        |    |
|-----------|--------|----|
| $r$       | $r$    | 3  |
|           | $\top$ | 17 |
| $\bar{r}$ | $r$    | 3  |
|           | $\top$ | 17 |

## Hypothèse 3 : $v$

|           |        |    |
|-----------|--------|----|
| $v$       | $r$    | 3  |
|           | $\top$ | 16 |
| $\bar{v}$ | $r$    | 3  |
|           | $\top$ | 18 |

On observe  $(\bar{c}r\bar{v}, \top)$

Quelle hypothèse retenir ?

# Distance entre distributions d'effets

## Estimation des probabilités d'effets

Après  $n$  observations parmi un ensemble d'effets  $\{e_1, \dots, e_q\}$  la probabilité estimée d'ocurrence d'un effet  $e_i$  est  $\frac{\#e_i}{n}$ .

## Distance entre estimations de distributions

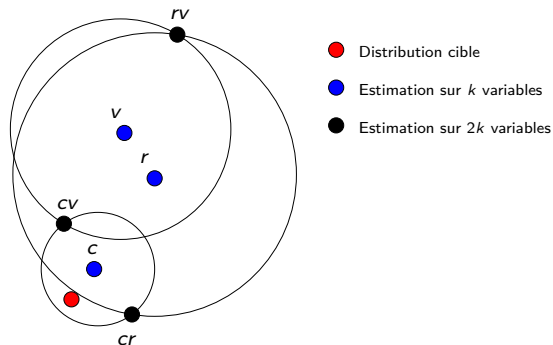
La distance  $L_1$  entre une estimation de probabilités sur  $n$  observations et une autre sur  $m$  observations est :

$$d = \sum_i \left| \frac{\#_1 e_i}{n} - \frac{\#_2 e_i}{m} \right|$$

# Retenir une hypothèse

## L'approche de SLF- $R_{\max}$ (Strehl *et al.* 2007)

Pour une action qui dépend de  $k$  variables, on classe les échantillons selon toutes les affectations possibles à  $2k$  variables.



- On garde l'hypothèse au centre du plus petit cercle.

## « Knows What It Knows »

Pour chaque hypothèse, on limite le nombre d'observations à  $m$  pour chaque prédicteur sur  $2k$  variables.

- un prédicteur est **connu** s'il a reçu  $m$  observations
- **inconnu** sinon.

## Exploration selon $R_{\max}$

- à chaque étape on planifie avec la meilleure hypothèse pour chaque état.
- tout état cohérent avec un prédicteur inconnu, à une  $Q$ -valeur importante → **encourage l'agent à apprendre ce qui n'est pas connu.**

## Efficacité

- Si  $k$  est une constante, le nombre d'étapes d'apprentissage est polynomial en  $m$ ,  $n$  et  $\binom{n}{2k}$ , avec une grande probabilité.

## Qualité

- L'erreur sur la fonction de valeur est polynomiale en  $m$ , avec une grande probabilité.

# Retour sur l'exemple du Loto

## Ambiguïté

Quand on est déjà riche, on ne peut pas observer directement si on a gagné ou non.

$$\text{Appliquer}(r, cr\bar{v}) = \text{Appliquer}(\top, cr\bar{v}) = cr\bar{v}$$

Les effets  $\top$ ,  $r$ ,  $c$ ,  $\bar{v}$ ,  $rc$ ,  $r\bar{v}$ ,  $c\bar{v}$  et  $cr\bar{v}$ , peuvent expliquer une transition de  $cr\bar{v}$  vers  $cr\bar{v}$ .

## Effets plausibles

Pour une transition de  $s$  vers  $s'$ , tous les effets  $e$  tels que :

$$s' \setminus s \subseteq e \subseteq s'$$

sont plausibles.

# Les hypothèses maintenues

Les observations que l'on classe sont des **intervalles d'effets**.

## Problème

La notion de distance entre distributions d'intervalles d'effets n'est pas correctement définie.

## Pour l'action de jouer au Loto

Une hypothèse regroupe les exemples qui satisfont  $c$ , c'est-à-dire aussi bien ceux qui satisfont  $c\bar{r}$  et  $cr$ . Dans le premier cas on peut identifier l'effet, mais pas dans le second.

# Observations ambiguës

Pour l'hypothèse qui classe les exemples qui satisfont  $c$

On observe les transitions  $(c\bar{r}v, crv)$ ,  $(c\bar{r}v, c\bar{r}v)$ ,  $(crv, crv)$ .

Transcription en intervalles d'effets :

- 1  $[r, crv] = \{r, cr, rv, crv\}$
- 2  $[\top, c\bar{r}v] = \{\top, c, \bar{r}, v, c\bar{r}, cv, \bar{r}v, c\bar{r}v\}$
- 3  $[\top, crv] = \{\top, c, r, v, cr, cv, rv, crv\}$

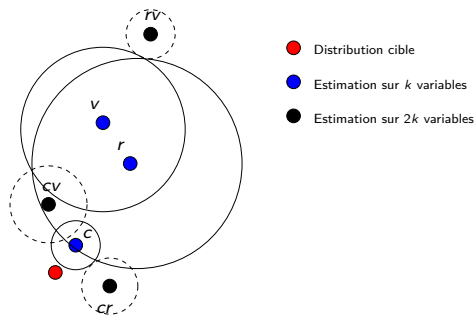
$$[\top, crv] \cap [r, crv] \neq \emptyset \text{ et } [\top, crv] \cap [\top, c\bar{r}v] \neq \emptyset$$

L'intervalle  $[\top, crv]$  peut donc correspondre à soit :

- un effet de  $[r, crv]$  ;
- un effet de  $[\top, c\bar{r}v]$  ;

# Distributions d'intervalles d'effets

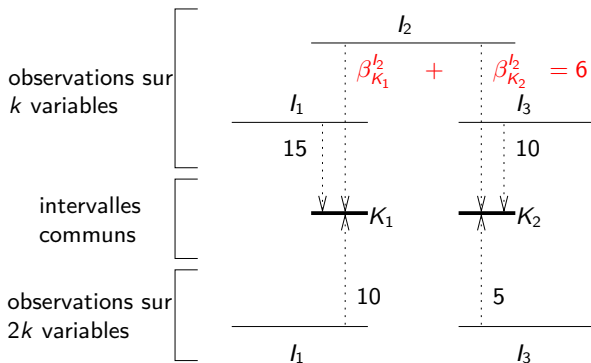
- L'incertitude sur la répartition des effets engendre plusieurs interprétations des observations
- On ne peut pas définir de manière unique la distance entre deux distributions.



## Une solution possible

Calculer une borne inférieure à la distance entre distributions.

# Distance $L_1$ entre distributions d'intervalles d'effets



## Calcul d'un encadrement de la distance

Avec un programme linéaire on peut optimiser les variables  $\beta$  pour calculer un encadrement de la distance.

$$d = \left| \frac{15 + \beta_{K_1}^{l_2}}{21} - \frac{10}{15} \right| + \left| \frac{10 + \beta_{K_2}^{l_2}}{21} - \frac{5}{15} \right|$$

# Nos contributions en quelques mots

Un algorithme d'apprentissage par renforcement

- Model based
- Avec un temps d'exploration polynomial à effets connus
- Efficace avec des effets inconnus (heuristique).

## Contribution principale

Calculer des distances entre distributions d'intervalles d'effets.

## Problèmes compacts en PSO

≈ 1200 étapes pour obtenir un comportement optimal (512 états - 9 actions)

## Problèmes non compacts en PSO

Nombre d'étapes similaires à  $R_{\max}$  (non factorisé)

## Problèmes compacts en PSO

≈ 1200 étapes pour obtenir un comportement optimal (512 états - 9 actions)

## Problèmes non compacts en PSO

Nombre d'étapes similaires à  $R_{\max}$  (non factorisé)

**Il faut passer à l'échelle.**

## En cours :

- Passer à l'échelle : utiliser un planificateur factorisé
- Utiliser la distance entre intervalles d'effets pour faire de l'induction d'arbres de décision pour les conditions.

## Quelques pistes

- Prendre en compte les états impossibles
- Apprendre des actions avec *aspects*.