



## MoCaNA

### UN AGENT DE NÉGOCIATION AUTOMATIQUE UTILISANT LA RECHERCHE ARBORESCENTE DE MONTE CARLO

Cédric Buron<sup>1,2</sup>, Zahia Guessoum<sup>1,3</sup>, Sylvain Ductor<sup>4</sup>, Olivier Roussel<sup>5</sup>

<sup>1</sup>LIP6, <sup>2</sup>Thales, <sup>3</sup>CRESTIC, <sup>4</sup>Université fédérale du Ceará, <sup>5</sup>Kyriba

[cedric.buron@thalesgroup.com](mailto:cedric.buron@thalesgroup.com)

11/10/2018

[www.thalesgroup.com](http://www.thalesgroup.com)

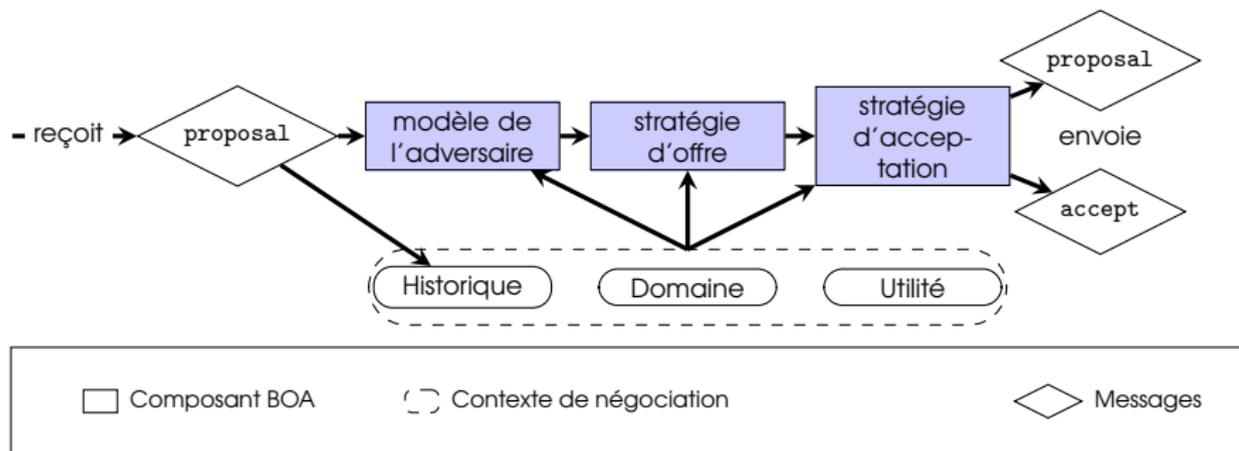


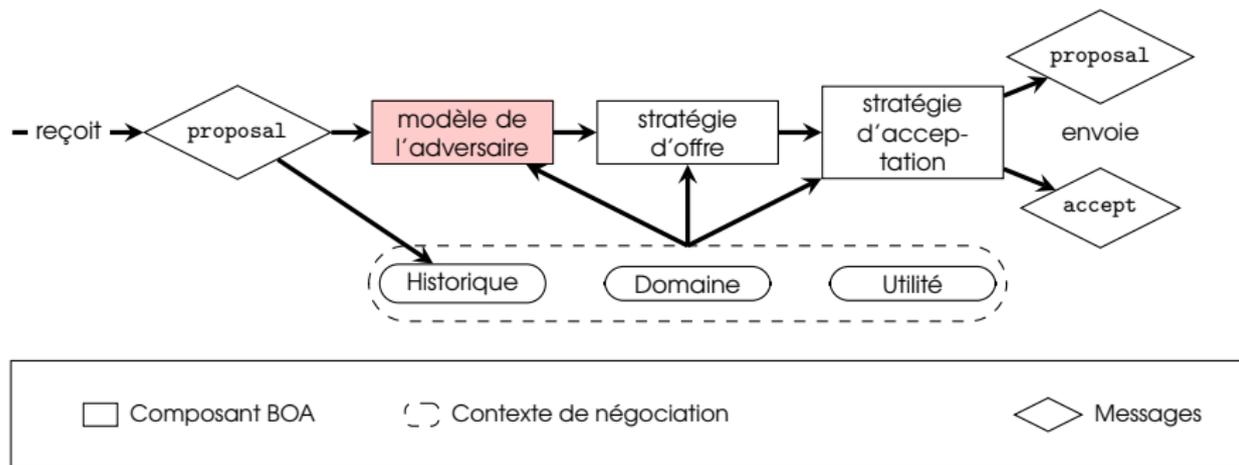
### Objectif : agent capable de négocier dans un contexte général :

- protocole : *bargaining*, avec ou sans borne
- biens : multi-attributs, attributs catégoriels et/ou continus

### Architecture BOA [1] :

- **B**idding Strategy ou stratégie d'offre
- **O**pponent modeling ou modélisation de l'adversaire
- **A**cceptance strategy ou stratégie d'acceptation





Utilisation d'apprentissage automatique pour modéliser certains aspects de l'adversaire [2]

5 éléments modélisés dans la littérature :

- | **stratégie d'offre**
- | **profil de préférence**
- | **stratégie d'acceptation**
- | **borne (temps ou nombre de tours)**
- | **prix de réserve**

Utilisation d'apprentissage automatique pour modéliser certains aspects de l'adversaire [2]  
5 éléments modélisés dans la littérature :

- 1 stratégie d'offre
- 1 profil de préférence
- 1 stratégie d'acceptation
- 1 borne (temps ou nombre de tours)
- 1 prix de réserve

### Processus gaussien : extension de la gaussienne multivariée

Soit  $(x_i, y_i)$ . On suppose :

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \sim \mathcal{N}(\mathbf{0}, K(x_i))$$

avec  $K$ , matrice de covariance représentant la proximité des tours entre eux, selon une fonction de covariance  $k$  :

$$K(x_i) = (k_{jk}) = (k(x_j, x_k))$$

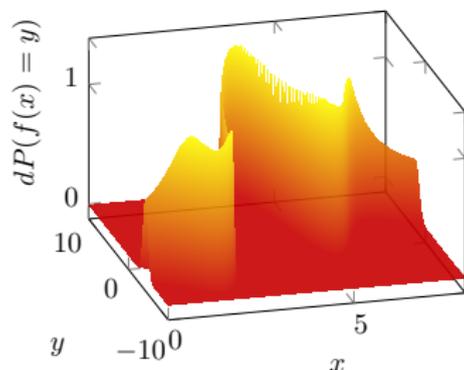
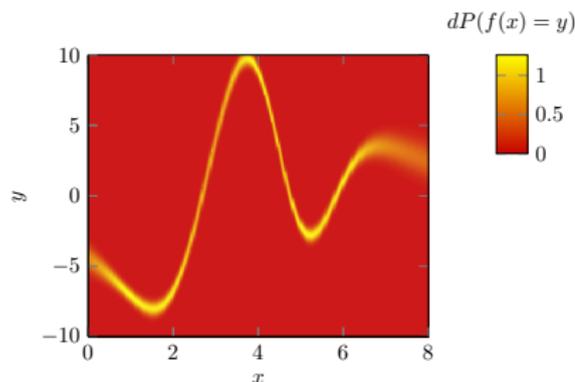
Alors la prédiction pour le tour  $x_*$  :

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \\ y_* \end{pmatrix} \sim \mathcal{N}(\mathbf{0}, K(x_i, x_*))$$

Théorème sur les gaussiennes multivariées :

$$\begin{cases} \bar{y}_* = K_* K^{-1} \mathbf{y} \\ \sigma_* = \text{var}(y_*) = K_{**} - K_* K^{-1} K_*^T \end{cases}$$

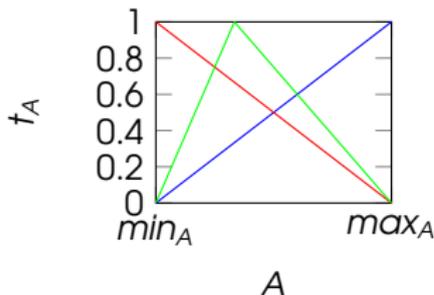
où  $K_* = (k(x_*, x_1), \dots, k(x_*, x_n))$  et  $K_{**} = k(x_*, x_*)$ .



### Généralisation aux attributs catégoriels

Hypothèses de l'apprentissage bayésien :

■ fonctions triangulaires  $t_i$  pour chaque attribut



■ un rang  $\pi_i$  pour l'attribut  $i$  et du poids  $w_i = 2 \frac{\pi_i}{n \cdot (n+1)}$

pour une proposition  $(x_1, \dots, x_n)$ , prédiction de l'hypothèse

$$h(b) = \sum_{1 \leq i \leq n} w_i \cdot t_i(x_i)$$

### Suppositions

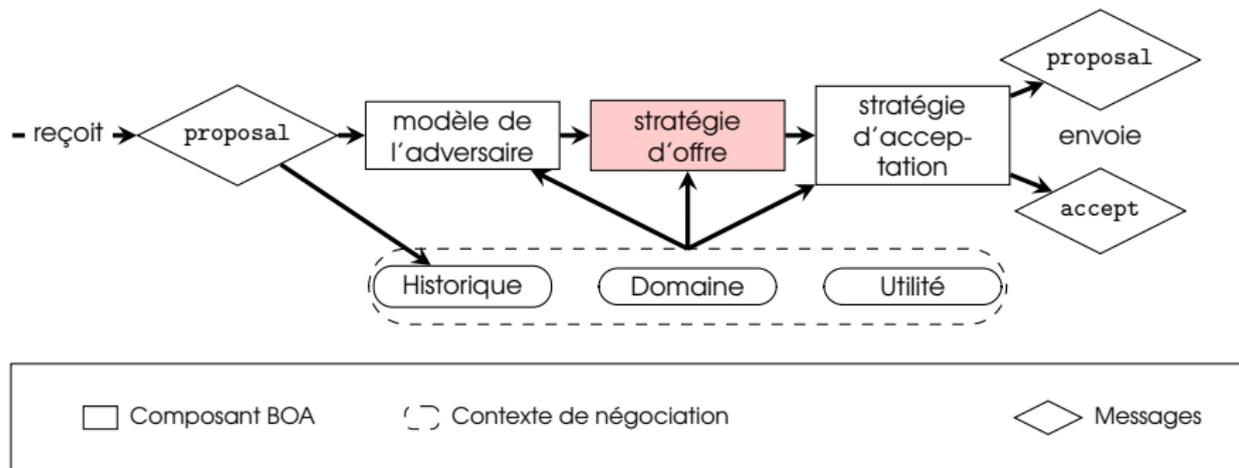
- **Probabilité de chaque hypothèse : gaussienne**
  - **Concessions monotones et régulières (proches de  $u^0 : b \mapsto 1 - \alpha_U \cdot \text{round}(b)$ , avec  $\alpha_U$  à déterminer)**
- Donc

$$P(b|h_j) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(u_j(b) - u^0(b))^2}{2\sigma^2}} \quad \text{et} \quad P(h_j|b) = \frac{P(h_j)P(b|h_j)}{\sum_{k \leq m} P(h_k)P(b|h_k)}$$

- Mise à jour de  $P(h_j)$  à chaque réception d'une proposition  $b$ .  
Modèle : somme pondérée des hypothèses

$$u = \sum_j P(h_j|b) \cdot h_j$$

- **Efficace pour les attributs numériques,**
- **Généralisation naturelle aux attributs catégoriels.**



**Sélection** de  $j$  maximisant [3] :

$$C_j = \frac{\bar{X}_j}{n_j + 1} + C \cdot n^{\alpha_s} \sqrt{\frac{\ln(n)}{n_j + 1}}$$

avec  $n$  le nombre total de simulations,  $n_j$  le nombre de simulations de  $j$  et  $\bar{X}_j$  le score moyen de  $j$ ,  $C$  et  $\alpha_s$  des paramètres du modèle.

**Expansion** d'un nouveau nœud ssi [3]

$$n_p^{\alpha_s} \geq n_c$$

avec  $n_p$  le nombre de simulation du parent et  $n_c$  le nombre de ses enfants.

**Simulation** basée sur la modélisation des stratégies d'offre et d'acceptation

**Rétropropagation** des scores ; score de l'agent et score modélisé de l'adversaire

### Simulation et rétropropagation reposant sur les 2 modèles de l'adversaire :

- ▶ modèle de la stratégie (régression de processus gaussien) [5]
- ▶ modèle de l'utilité (apprentissage bayésien) [4]

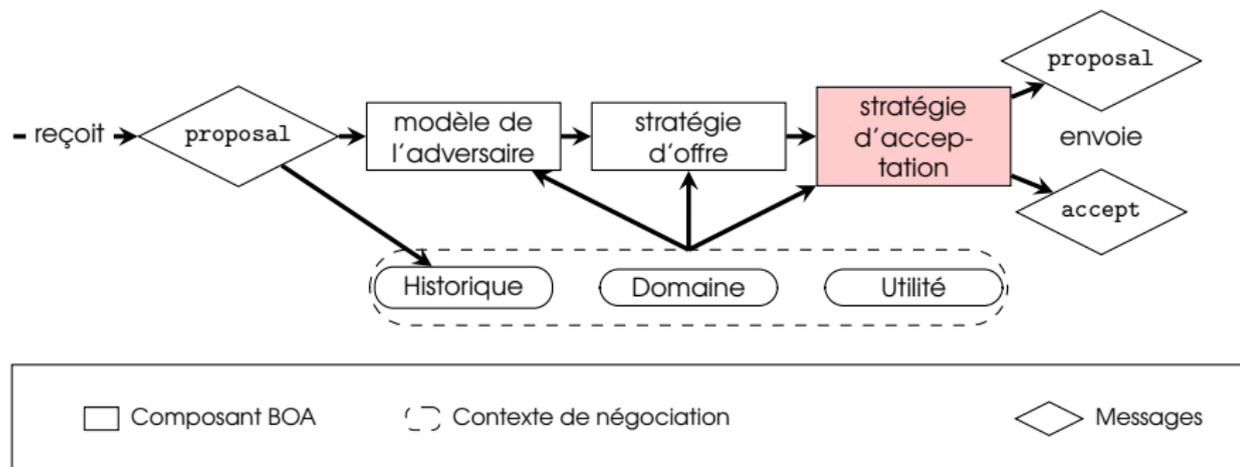
### Deux variantes d'élagage

#### 1. selon une valeur fixe :

- suppression des nœuds dont l'utilité est au-dessous d'un seuil,
- valeur déterminée en fonction de l'utilité attendue,
- ici, valeur seuil de 0.8.

#### 2. selon les offres de l'adversaire :

- suppression des nœuds dont l'utilité est pire que la meilleure proposition adverse,
- pas de nécessité de déterminer une utilité attendu *a priori*.



Trois types de stratégies :

### ■ stratégies myopes

- accepter si proposition adverse meilleure que dernière proposition de l'agent
- accepter si proposition adverse meilleure que prochaine proposition de l'agent
- accepter si proposition adverse meilleure que meilleure proposition de l'agent
- accepter si proposition adverse meilleure qu'un seuil prédéfini

### ■ mélange de stratégies myopes, selon la distance à la borne

### ■ stratégie « optimale »

- modélisation la stratégie adverse
- calcul du tour auquel la proposition adverse sera la meilleure
- mise à jour chaque tour
- accepter ssi le tour est optimal et la proposition est acceptable

Trois types de stratégies :

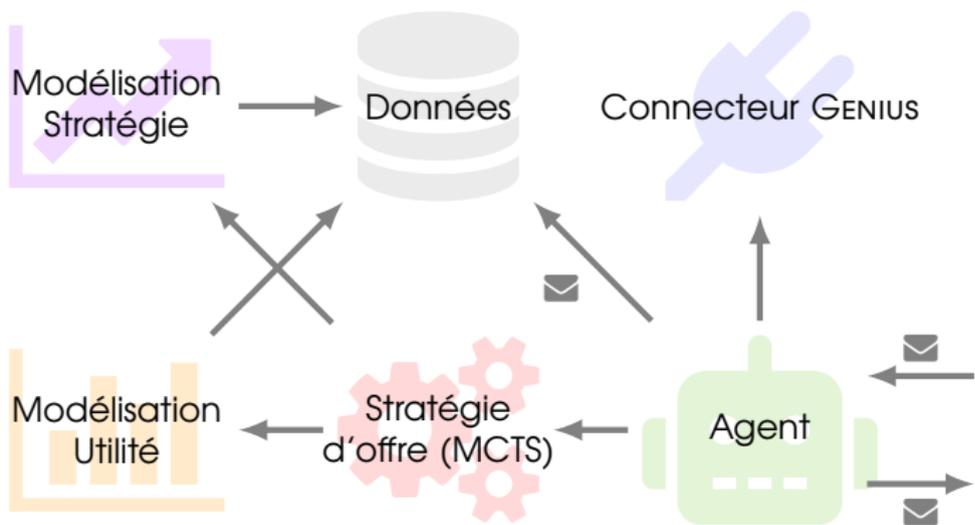
### ■ stratégies myopes

- accepter si proposition adverse meilleure que dernière proposition de l'agent
- accepter si proposition adverse meilleure que prochaine proposition de l'agent
- accepter si proposition adverse meilleure que meilleure proposition de l'agent
- accepter si proposition adverse meilleure qu'un seuil prédéfini

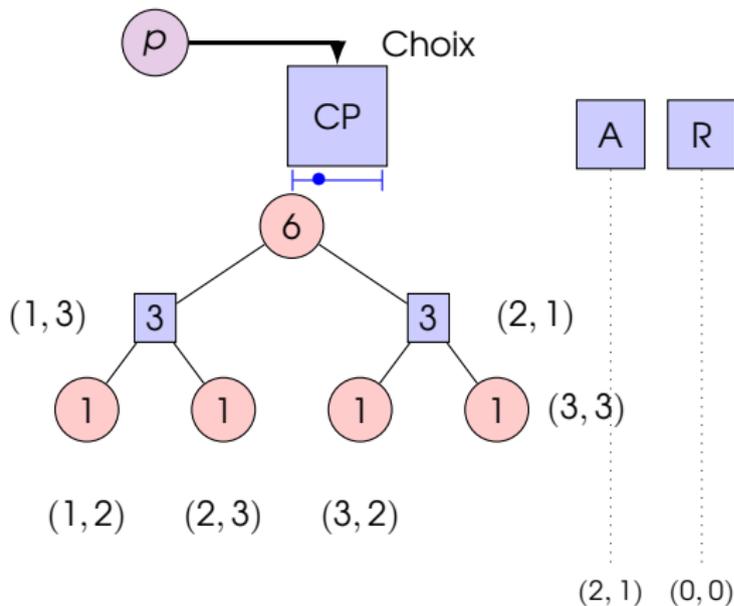
### ■ mélange de stratégies myopes, selon la distance à la borne

### ■ stratégie « optimale »

- modélisation la stratégie adverse
- calcul du tour auquel la proposition adverse sera la meilleure
- mise à jour chaque tour
- accepter ssi le tour est optimal et la proposition est acceptable



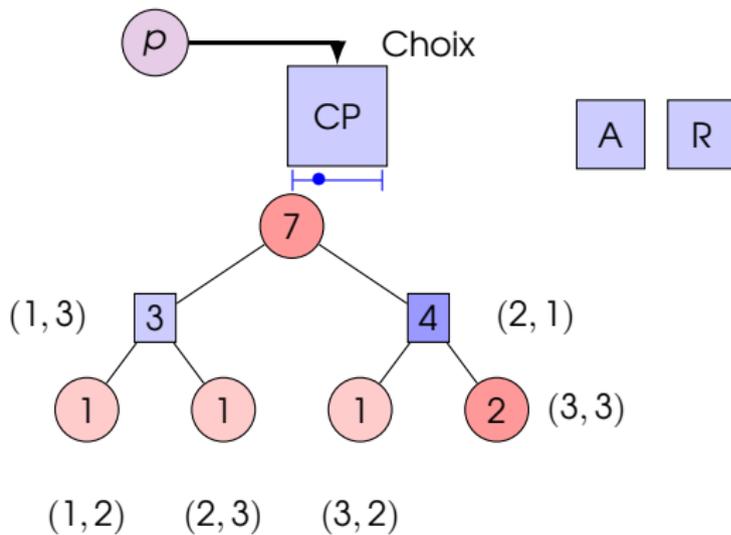
## Exemple



## Exemple



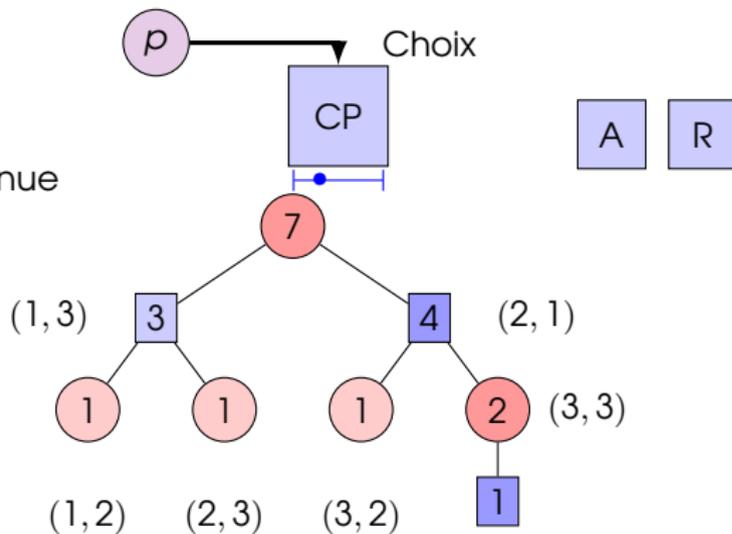
Sélection  
Critère :  $\max^n$



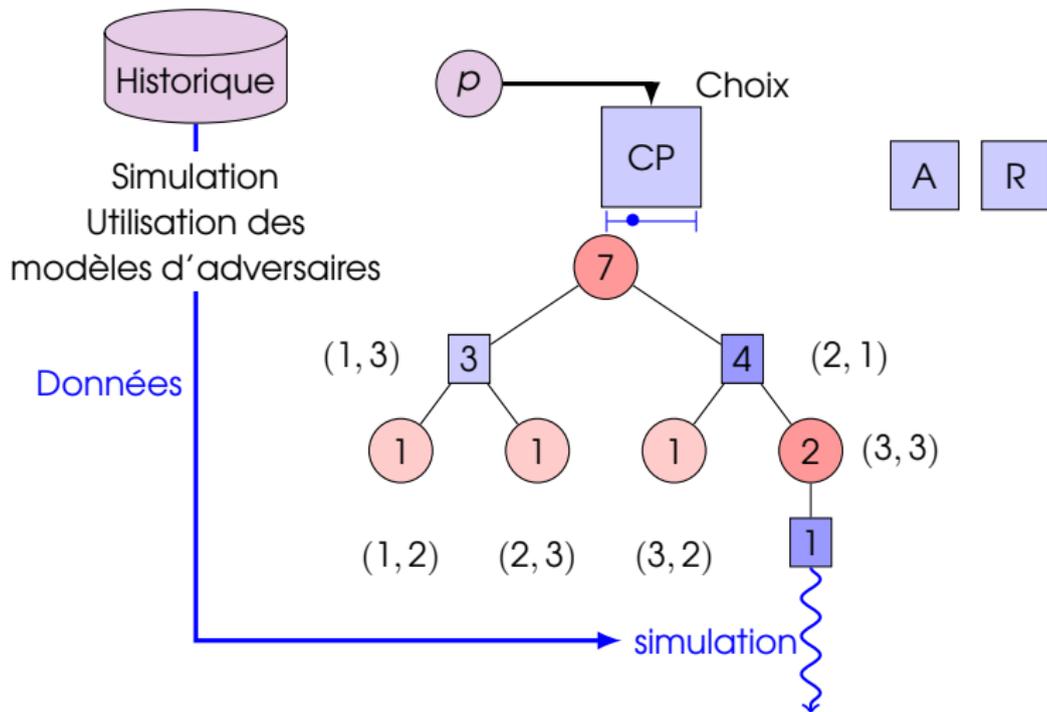
## Exemple



Extension  
Proposition : continue



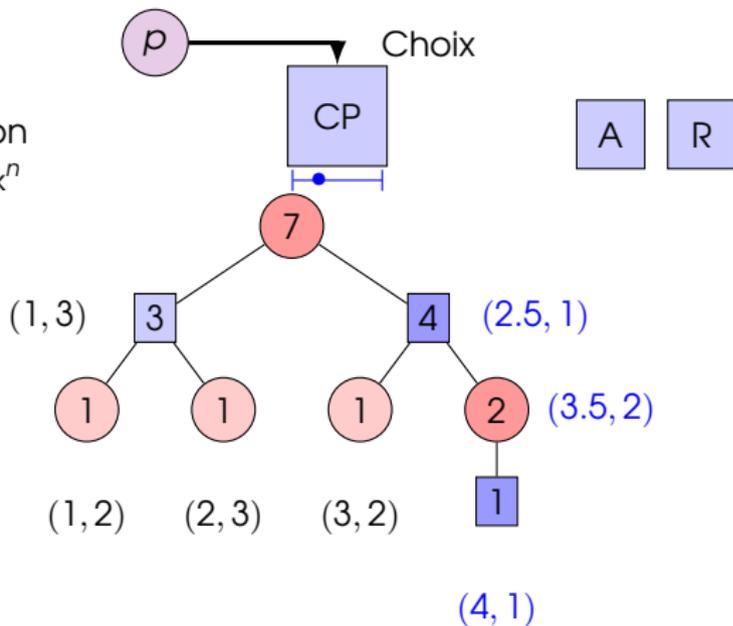
## Exemple



## Exemple



Rétropropagation  
mise à jour :  $\max^n$



## utilisation des agents ANAC 2014

- présence d'une échéance,
- domaines très larges mais non continus,
- 2 profils de préférence (somme non nulle),
- fonctions d'utilité non linéaires,
- 10 attributs,

## implémentation d'un *RandomWalker*

- pas d'échéance,
- même domaine, mêmes profils d'utilité.

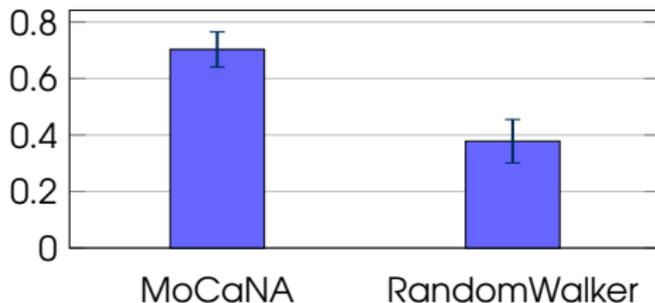
## lancé sous GENIUS, 3min/tour,

## moyenné sur 20 négociation (10/profil),

## Évaluation : résultats sans élagage

Adversaire	Score adversaire	Score MoCaNA	Taux d'accord
AgentM	0.882 ( $\pm 0.069$ )	0.655 ( $\pm 0.076$ )	0.35
DoNA	1.000 ( $\pm 0.000$ )	0.572 ( $\pm 0.000$ )	0.05
Gangster	0.814 ( $\pm 0.130$ )	0.637 ( $\pm 0.142$ )	0.45
Whale	0.779 ( $\pm 0.164$ )	0.697 ( $\pm 0.113$ )	0.45
Group2	0.689 ( $\pm 0.211$ )	0.822 ( $\pm 0.194$ )	0.45
kGAgent	0.977 ( $\pm 0.085$ )	0.554 ( $\pm 0.047$ )	0.20
AgentYk	0.806 ( $\pm 0.174$ )	0.752 ( $\pm 0.154$ )	0.15
BraveCat	0.622 ( $\pm 0.170$ )	0.749 ( $\pm 0.101$ )	0.95

MoCaNA sans élagage VS ANAC 2014



MoCaNA sans élagage VS RandomWalker

## Évaluation : résultats avec élagage variable/fixe

Adversaire	Score adversaire	Score MoCaNA	Taux d'accord
AgentM	0.770 ( $\pm 0.113$ )	0.656 ( $\pm 0.077$ )	0.10
DoNA	0.857 ( $\pm 0.005$ )	0.547 ( $\pm 0.058$ )	0.10
Gangster	0.563 ( $\pm 0.221$ )	0.702 ( $\pm 0.162$ )	0.55
Whale	0.639 ( $\pm 0.085$ )	0.768 ( $\pm 0.097$ )	0.25
Group2	0.534 ( $\pm 0.075$ )	0.786 ( $\pm 0.098$ )	0.45
kGAgent	1 ( $\pm 0.0$ )	0.485 ( $\pm 0.0$ )	0.05
AgentYk	0.771 ( $\pm 0.128$ )	0.637 ( $\pm 0.106$ )	0.15
BraveCat	0.519 ( $\pm 0.232$ )	0.662 ( $\pm 0.121$ )	1.0

Adversaire	Score adversaire	Score MoCaNA	Taux d'accord
AgentM	0.699 ( $\pm 0.078$ )	0.847 ( $\pm 0.016$ )	0.15
DoNA	N/A	N/A	0
Gangster	0.534 ( $\pm 0.193$ )	0.852 ( $\pm 0.049$ )	0.55
Whale	0.708 ( $\pm 0.064$ )	0.832 ( $\pm 0.020$ )	0.45
Group2	0.530 ( $\pm 0.061$ )	0.870 ( $\pm 0.039$ )	0.45
kGAgent	N/A	N/A	0
AgentYk	0.712 ( $\pm 0.074$ )	0.841 ( $\pm 0.025$ )	0.10
BraveCat	0.487 ( $\pm 0.195$ )	0.855 ( $\pm 0.066$ )	0.95

### Agent très efficace dans le cas sans échéance

### Cas avec échéance :

- ▶ baisse du taux d'accord
  - pas de « Time Pressure »
  - diminution en cas d'élagage
- ▶ score en cas d'accord :
  - élagage fixe : meilleur que tous les finalistes avec lesquels il trouve un accord
  - élagage variable : meilleur que la moitié des finalistes
  - sans élagage : meilleur que BraveCat et Group2 uniquement

## Contributions

- capable de négocier dans un contexte très général (protocole, attributs)
- très efficace dans le contexte sans borne
- efficace dans un contexte avec borne

## Perspectives :

- autres améliorations de MCTS :
  - AMAF
  - RAVE
  - ...
- implémentation GPGPU
- adaptation d'autres agents au contexte sans borne et comparaison



**Tim BAARSLAG.** "Exploring the Strategy Space of Negotiating Agents : A Framework for Bidding, Learning and Accepting in Automated Negotiation". In : *Cham : Springer International Publishing*, 2016. Chap. A Component-Based Architecture to Explore the Space of Negotiation Strategie, p. 53-69.



**Tim BAARSLAG et al.** "Learning about the opponent in automated bilateral negotiation : a comprehensive survey of opponent modeling techniques". In : *Autonomous Agents and Multi-Agent Systems 20.1* (2015), p. 1-50.



**Adrien Couëroux.** "Monte Carlo Tree Search for Continuous and Stochastic Sequential Decision Making Problems". *Thèse de doct. Université Paris XI*, 2013.



**Koen HINDRIKS et Dmytro TYKHONOV.** "Opponent Modelling in Automated Multi-issue Negotiation Using Bayesian Learning". In : *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems*. T. 1. International Foundation for Autonomous Agents et Multiagent Systems, 2008, p. 331-338.



**Colin R WILLIAMS et al.** "Using Gaussian Processes to Optimise Concession in Complex Negotiations Against Unknown Opponents". In : *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume One*. IJCAI'11. AAAI Press, 2011, p. 432-438.