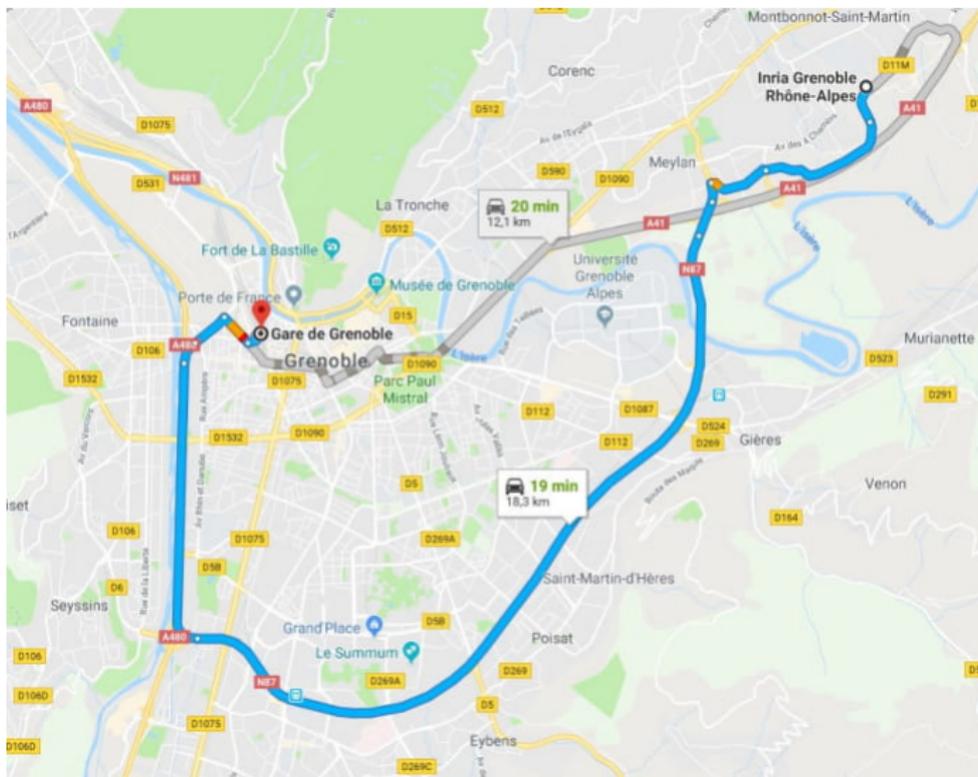


Analyse de la dynamique de meilleure réponse dans les jeux de potentiel

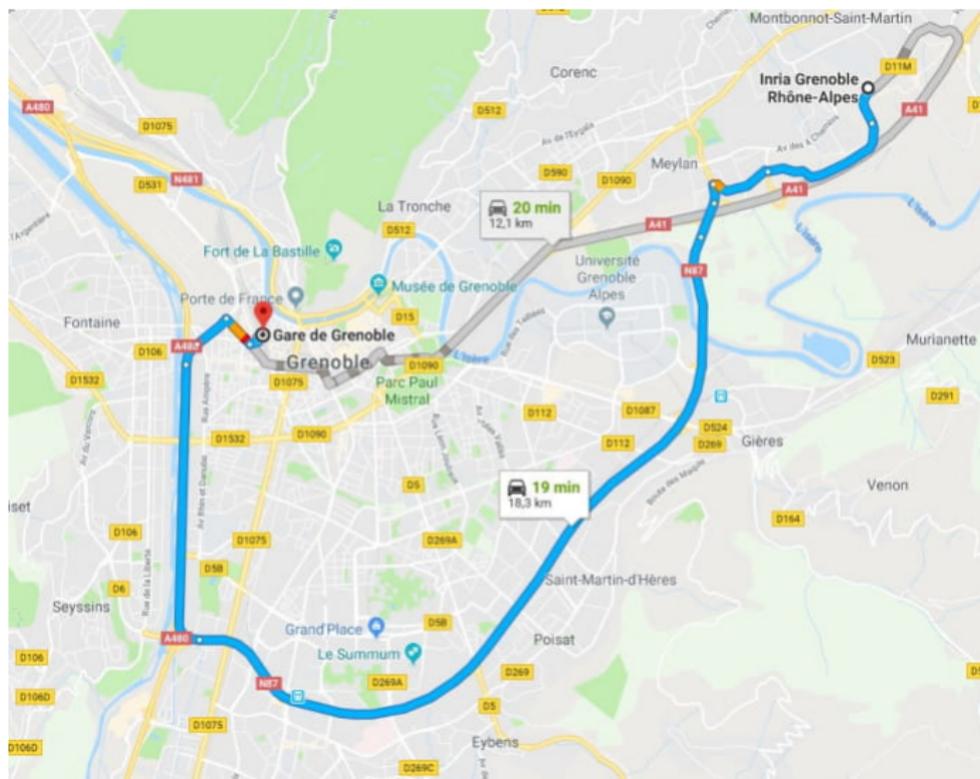
Stéphane Durand
encadré par Federica Garin et Bruno Gaujal



Un exemple d'application de la théorie des jeux



Un exemple d'application de la théorie des jeux



→ Jeu à n joueurs, 2 stratégies par joueurs

Théorie des jeux

Un **Jeu** est un ensemble de n **joueurs** chacun choisissant une **stratégie** et muni d'une **fonction d'utilité** u_i . On appelle **état** le vecteur des stratégies des joueurs.

Exemple classique : dilemme du prisonnier

u_1	C	T
T	5	1
C	3	0

u_2	C	T
T	0	1
C	3	5

Utilités des joueurs.

Deux joueurs, ayant deux stratégies chacun (Coopérer ou Trahir). Les tableaux montrent leurs utilités respectives.

Jeux de potentiel

Un jeu est **de potentiel** s'il admet une **fonction de potentiel** V telle que si un joueur change seul de choix, la variation de potentiel est la variation d'utilité de ce joueur.

Le dilemme du prisonnier est un jeu de potentiel

u_1	C	T		u_2	C	T		V	C	T
T	5	1		T	0	1	→	T	2	3
C	3	0		C	3	5		C	0	2

Beaucoup de jeux usuels sont des jeux de potentiel, en particulier, tous les jeux de congestions sont de potentiel ([Monderer, Shapley 1996]).

À partir de maintenant, on décrit les jeux par leur potentiels.

Jeux de potentiel

Un jeu est **de potentiel** s'il admet une **fonction de potentiel** V telle que si un joueur change seul de choix, la variation de potentiel est la variation d'utilité de ce joueur.

Le dilemme du prisonnier est un jeu de potentiel

u_1	C	T		u_2	C	T		V	C	T
T	5	1		T	0	1	→	T	2	3
C	3	0		C	3	5		C	0	2

Beaucoup de jeux usuels sont des jeux de potentiel, en particulier, tous les jeux de congestions sont de potentiel ([Monderer, Shapley 1996]).

À partir de maintenant, on décrit les jeux par leur potentiels.

Jeux de potentiel

Un jeu est **de potentiel** s'il admet une **fonction de potentiel** V telle que si un joueur change seul de choix, la variation de potentiel est la variation d'utilité de ce joueur.

Le dilemme du prisonnier est un jeu de potentiel

u_1	C	T		u_2	C	T		V	C	T
T	5	1		T	0	1	→	T	2	3
C	3	0		C	3	5		C	0	2

Beaucoup de jeux usuels sont des jeux de potentiel, en particulier, tous les jeux de congestions sont de potentiel ([Monderer, Shapley 1996]).

À partir de maintenant, on décrit les jeux par leur potentiels.

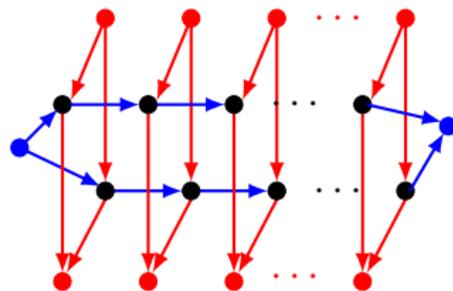
Exemple: Jeux de routage

Jeu de routage:

- Joueurs : flots
- Stratégies : choix du chemin
- Utilités: opposé de la somme des latences des nœuds traversés, croissantes selon le nombre de joueurs traversant le nœud.

Ces jeux sont toujours de potentiel.

Exemple :



Fonction de meilleure réponse et Équilibres de Nash

Dans l'état \mathbf{x} , la **meilleure réponse** $MR_k(\mathbf{x})$ d'un joueur k est l'ensemble de ses choix maximisant son utilité (\Leftrightarrow maximisant le potentiel).

V:

8	2	1	◇9
1	2	◇7	4
③	①	0	③
1	④	2	8

Un **équilibre de Nash** est un point fixe de cette fonction pour tous les joueurs. C'est également un maximum local du potentiel.

Dynamique de meilleure réponse (DMR)

La dynamique de meilleure réponse est la famille des algorithmes explorant les états du jeu via la fonction de meilleure réponse.

8	2	1	9
1	2	7	4
3	1	0	3
1	4	2	8

The diagram shows a 4x4 matrix of payoffs. A path of best responses is indicated by arrows: from (3,2) to (3,1), from (3,1) to (4,1), from (4,1) to (4,3), and from (4,3) to (1,4). The cell (1,4) with payoff 9 is circled in blue and has a self-loop arrow, indicating it is a Nash equilibrium.

Ce parcours s'arrête si et seulement si le dernier état est un équilibre de Nash.

Étude de la dynamique de meilleure réponse:

- Complexité moyenne dans le cas séquentiel ([SAGT 2016](#))
- Avec jeu simultané
 - Condition de convergence lors de jeux simultanés ([NetGCoop 2014](#))
 - Complexité moyenne dans le cas distribué ([IFIP Performance 2018](#) et [Performance Evaluation](#))
- Avec des joueurs indépendants ([ECC 2018](#))

Travail annexe sur la simulation parfaite:

- Algorithme cubique déterminant les queues stables et instables d'un réseau de Jackson ([Queuing Systems 2015](#))
- Extension de la méthode d'enveloppe pour grilles avec arêtes interdites ([QUEST 2014](#))

- 1 Introduction
- 2 Cas séquentiel
- 3 Avec jeu simultané
- 4 Profiter d'une structure
- 5 Conclusion

- On se place dans l'espace des jeux de potentiel, chaque potentiel tiré indépendamment uniformément sur $[0, 1]$.
- Ceci est équivalent à tirer une fonction de potentiel uniformément parmi toutes les fonctions possibles.
- La **complexité moyenne** est la moyenne des temps d'exécution de l'algorithme sur chaque jeu.

Théorème

Pour un jeu de n joueurs et a stratégies chacun, le nombre d'étapes moyen de l'algorithme suivant une séquence round robin est

$$\mathbb{E}T_{DMR} = e^\gamma n + o(n) \text{ lorsque } n \rightarrow \infty \text{ et } a \text{ quelconque}$$

avec γ la constante d'Euler ($e^\gamma \simeq 1.78$).

- Contraste : Linéaire contre le pire cas exponentiel
- Nombre d'étapes indépendant de a
- Constante faible, joueurs jouent moins de 2 fois en moyenne.

Étapes de la preuve

- ① On construit une approximation markovienne de l'évolution de l'algorithme.
- ② On compare le temps d'exécution de l'algorithme avec celui de l'approximation.
- ③ On analyse la complexité de l'approximation.

Cette approche est précise et donne la constante.

(1) Approximation sans mémoire

Approximation sans mémoire (Intersection Free Approximation : IFA)

- Valeurs de potentiels choisies à la volée
- Valeurs oubliées après usage. On ne garde que l'état et le potentiel actuel.

On appelle **intersections** les états dont le potentiel est lu à nouveau. L'algorithme et l'approximation diffèrent uniquement lors des intersections.

Intersections

Comportement réel

	1		

Sous IFA

	1		

Intersections

Comportement réel

0	1	5	2

Sous IFA

0	1	5	2

Intersections

Comportement réel

		6	
		2	
0	1	5	2
		4	

Sous IFA

		6	
		2	
0	1	5	2
		4	

Intersections

Comportement réel

7	4	6	0
		2	
0	1	5	2
		4	

Sous IFA

7	4	6	0
		2	
		5	
		4	

Intersections

Comportement réel

7	4	6	0
8		2	
0	1	5	2
9		4	

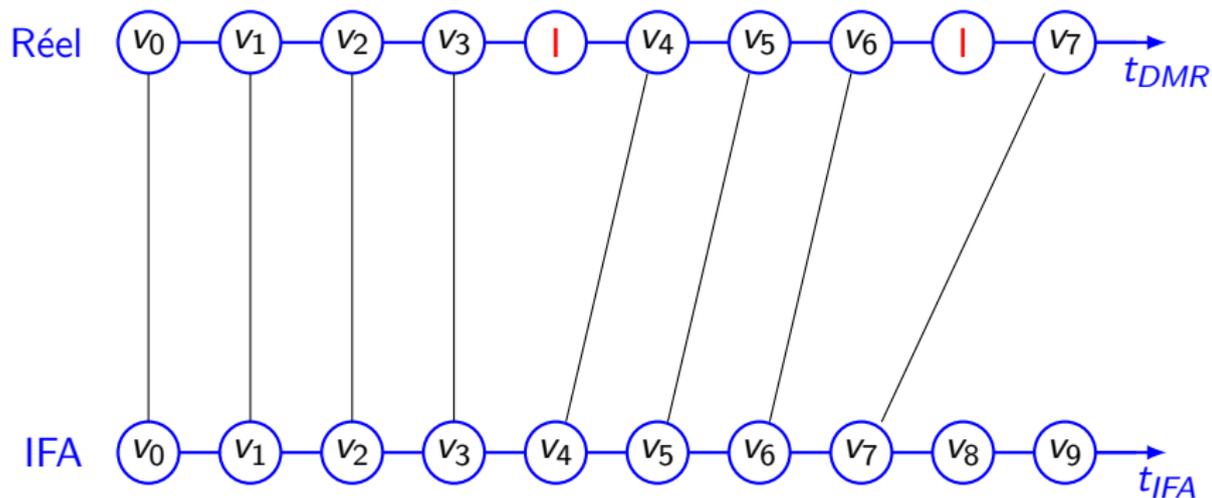
La valeur 0 est connue, la nouvelle valeur 9 est appliquée au tirage suivant.

Sous IFA

7	4	6	0
8			
9			
3			

La valeur 0 associée à l'état (1,2) a été oubliée au profit du nouveau tirage

Couplage des potentiels



Intersections

Comportement réel

7	4	6	0
8		2	
0	1	5	2
9	3	4	1

Sous IFA

7			
8			
9	3	1	6
3			

L'équilibre de Nash est un état différent dans les deux jeux, mais correspond au même potentiel tiré.

Le potentiel tiré correspondant à l'EN atteint par IFA est toujours un EN pour le jeu réel.

Conséquence: Comparaison de IFA et DMR

Lemme

Pour toute paire de jeu associés :

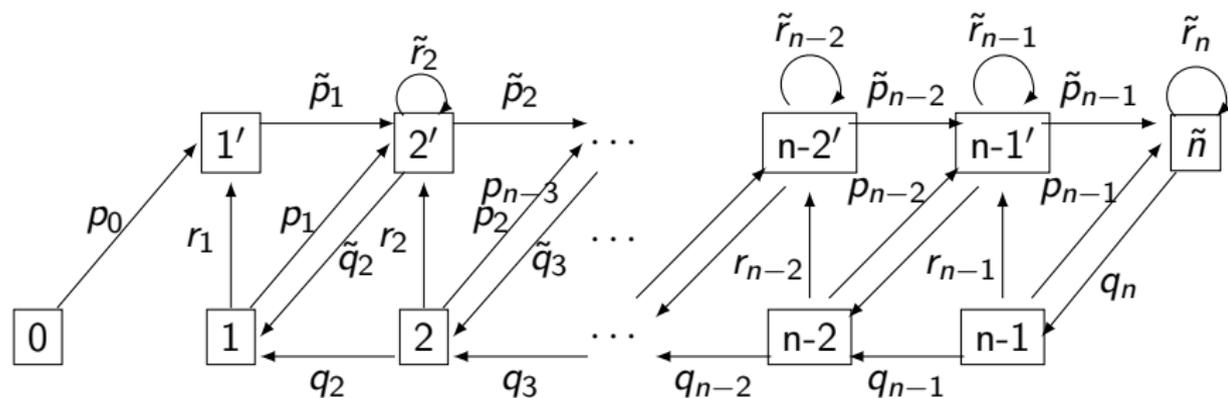
$$t_{DMR} \leq t_{IFA} + l$$

Donc

$$T_{DMR} \leq T_{IFA} + \frac{l}{a}$$

(2) Nombre d'intersections

Le nombre d'intersection est obtenu par la résolution d'une chaîne de Markov, en utilisant des propriétés du temps d'atteinte issues du livre de [Aldous and Fill, 2002]

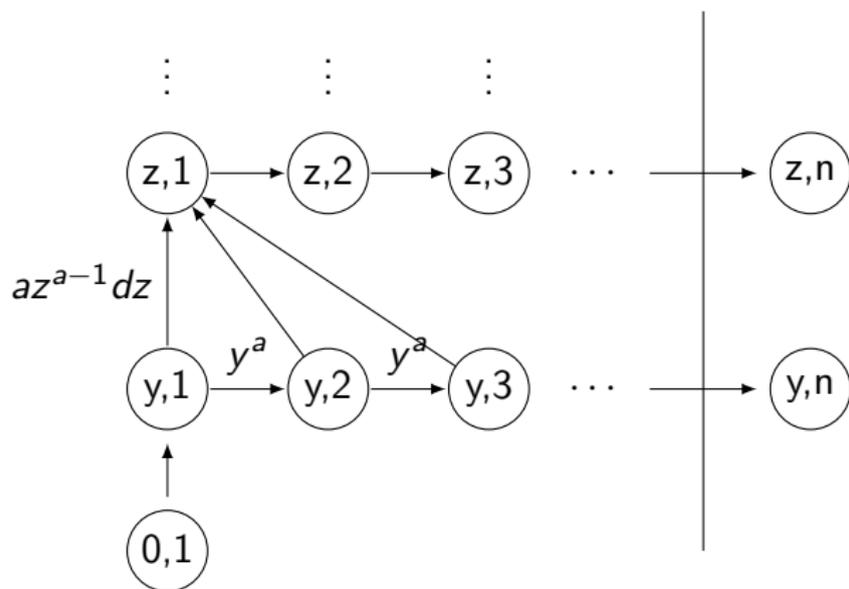


Chaîne de la distance de Hamming entre un état fixe et l'état courant.

Les intersections ne sont possibles qu'aux distances 1 et 2.

(3) Analyse de IFA

Les états sont les couples (potentiel, nb joueurs satisfaits).



Le temps moyen d'atteinte des états $(*, n)$ satisfait un système différentiel qu'on résout explicitement.

- 1 Introduction
- 2 Cas séquentiel
- 3 Avec jeu simultané**
- 4 Profiter d'une structure
- 5 Conclusion

Conséquence de jeu simultané

Deux joueurs actifs en même temps interfèrent et risquent une diminution de potentiel

1	9	6	4
6	3	5	0
2	5	7	2
1	2	3	7

Conséquence de jeu simultané

Ou même un cycle, risquant des trajectoires infinies

1	9	6	4
6	3	5	0
2	5	7	2
1	2	3	7

Séparabilité

Une famille d'ensemble est **séparable** sur un ensemble de joueur si

- La réunion des ensembles couvre tous les joueurs.
- La famille contient un singleton.
- En enlevant le joueur correspondant, la famille reste séparable.

Exemples :

- Une famille contenant tous les singletons est séparable. (on retrouve le premier résultat de [Alos-Ferrer, Netzer 2010])
- $\{1\}, \{1, 2\}, \{1, 3\}, \{2, 4\}, \{1, 3, 4\}$ est séparable sur $\{1, 2, 3, 4\}$.

Séparabilité

Une famille d'ensemble est **séparable** sur un ensemble de joueur si

- La réunion des ensembles couvre tous les joueurs.
- La famille contient un singleton.
- En enlevant le joueur correspondant, la famille reste séparable.

Exemples :

- Une famille contenant tous les singletons est séparable. (on retrouve le premier résultat de [Alos-Ferrer, Netzer 2010])
- $\{1\}, \{1, 2\}, \{1, 3\}, \{2, 4\}, \{1, 3, 4\}$ est séparable sur $\{1, 2, 3, 4\}$.

Séparabilité

Une famille d'ensemble est **séparable** sur un ensemble de joueur si

- La réunion des ensembles couvre tous les joueurs.
- La famille contient un singleton.
- En enlevant le joueur correspondant, la famille reste séparable.

Exemples :

- Une famille contenant tous les singletons est séparable. (on retrouve le premier résultat de [Alos-Ferrer, Netzer 2010])
- $\{1\}, \{1, 2\}, \{1, 3\}, \{2, 4\}, \{1, 3, 4\}$ est séparable sur $\{1, 2, 3, 4\}$.

Séparabilité

Une famille d'ensemble est **séparable** sur un ensemble de joueur si

- La réunion des ensembles couvre tous les joueurs.
- La famille contient un singleton.
- En enlevant le joueur correspondant, la famille reste séparable.

Exemples :

- Une famille contenant tous les singletons est séparable. (on retrouve le premier résultat de [Alos-Ferrer, Netzer 2010])
- $\{1\}, \{1, 2\}, \{1, 3\}, \{2, 4\}, \{1, 3, 4\}$ est séparable sur $\{1, 2, 3, 4\}$.

Séparabilité

Une famille d'ensemble est **séparable** sur un ensemble de joueur si

- La réunion des ensembles couvre tous les joueurs.
- La famille contient un singleton.
- En enlevant le joueur correspondant, la famille reste séparable.

Exemples :

- Une famille contenant tous les singletons est séparable. (on retrouve le premier résultat de [Alos-Ferrer, Netzer 2010])
- $\{1\}, \{1, 2\}, \{1, 3\}, \{2, 4\}, \{1, 3, 4\}$ est séparable sur $\{1, 2, 3, 4\}$.

Condition de convergence

Un processus de révision est séparable ssi:

- À chaque étape, tirage indépendant d'un ensemble de joueurs.
- La famille des ensembles de probabilité non nulle est séparable.

Théorème convergence presque sûre

Processus de révision séparable.

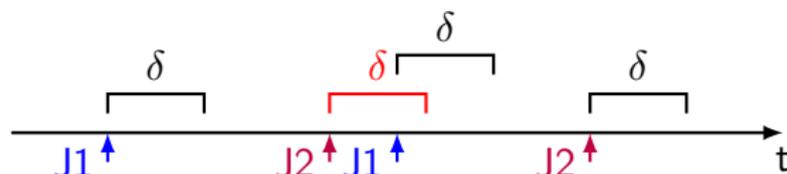


Dynamique converge presque sûrement sur tout jeu

Algorithme distribué avec horloge de Poisson

- L'algorithme de meilleure réponse est similaire au cas round robin.
- Chaque joueur se réveille selon un processus de Poisson de taux $\frac{\lambda}{n}$
- La superposition des activations forme un processus de Poisson de taux λ

Jeu simultané et collisions



- La fonction meilleure réponse prend un temps non nul δ de calcul et mise en place.
- Deux joueurs jouent trop près \rightarrow ils interfèrent \rightarrow état de potentiel inconnu.
On appelle cela une **collision**, équivalent à faire jouer les joueurs ensemble.
- Avec taux λ : $P(\text{collision}) = 1 - e^{-\lambda\delta}$.
- Le processus de révision correspondant est homogène et contient tous les singletons \Rightarrow séparable.

Résultat, en négligeant les collisions

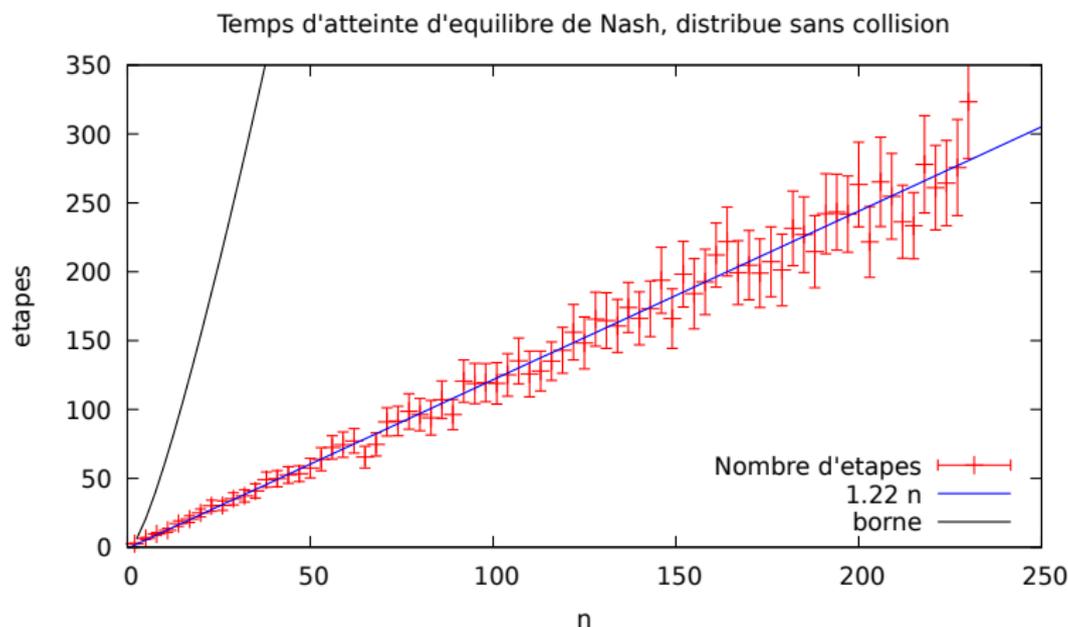
Si $\delta \ll \frac{1}{\lambda}$, on peut négliger les collisions.

Théorème Complexité moyenne sans collision

Le nombre moyen d'étapes avant convergence est majoré par

$$2e^{\gamma} n \log n + O(n)$$

Temps d'atteinte simulé



Temps d'atteinte d'un équilibre de Nash, sans collision. ($a = 20$
1000 répétitions par points, n variable)

Le temps total simulé est donc $1.22n + n \log n < 2e^\gamma n \log n$

Résultat, avec collision

Le temps moyen d'exécution avec collision et un taux λ est minoré par $\frac{nH_n}{\lambda}$ et majoré par

$$\frac{1}{\lambda} \left((n^{1-p} H_{n-1} + 1) G_{n,p} + \frac{n^{1-p} + C_1 + \frac{C_2}{n} + C_2 \frac{G(p,n)}{n(1-p)}}{1-p} G_{n,p} \right).$$

avec $G_{n,p} = \left(\frac{e^{pH_{n-1}}}{p} - \frac{1-p}{p} \right)$ majorable par $(n-1)e$.

$H_n \simeq \log n$ est le n -ème nombre harmonique.

$p = 1 - e^{-\delta\lambda}$ est la probabilité de collision.

C_1 et C_2 sont des constantes.

On voit le terme principal en $n \log n$ et le terme en $\frac{1}{1-p}$ causé par les collisions.

Résultat, avec λ choisi

Optimisation du taux

En prenant $\hat{\lambda} = \frac{\log(\log(n)) - \log \log \log(n)}{\delta}$ (et donc $\hat{p} = 1 - \frac{\log(\log(n))}{\log(n)}$),
Le temps d'exécution moyen vérifie

$$\frac{\delta n \log(n)}{\log(\log(n))} \leq \mathbb{E} T \leq e^{\gamma} \frac{\delta n \log(n)}{\log(\log(n))}$$

avec γ la constante d'Euler.

- La probabilité de collision \hat{p} tend (lentement) vers 1.
- L'augmentation du taux λ permet d'améliorer le $n \log n$.

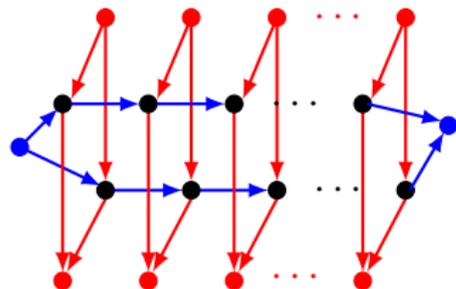
- 1 Introduction
- 2 Cas séquentiel
- 3 Avec jeu simultané
- 4 Profiter d'une structure**
- 5 Conclusion

Joueurs indépendants

- Un joueur est **indépendant** d'un autre si leurs meilleures réponses commutent.
- En particulier, ceci est vrai si les utilités d'un joueur ne dépendent pas des stratégies d'un autre.
- Les collisions entre joueurs indépendants sont sans effet car équivalentes à un jeu successif.

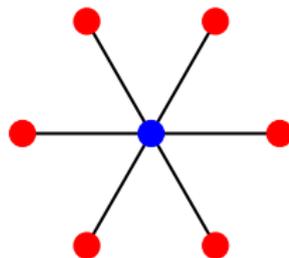
Exemple

On reprend l'exemple de jeu de routage :



Les joueurs rouges sont indépendant deux à deux.

On appelle **Graphe d'interaction** le graphe des joueurs non indépendants.



Graphe d'interaction de l'exemple.

Cas round robin : Couleur

On considère la procédure

- Pré-calcul: on obtient une coloration du graphe des joueurs.
- On lance la dynamique de manière round robin sur les couleurs

Une étape de cette procédure est l'exécution parallèle de la fonction meilleure réponse par un grand nombre de joueurs indépendants.

Espérance du nombre d'étapes: $e^\gamma \sigma$ avec σ le nombre chromatique obtenu pour le graphe lors du pré-calcul.

Cas réseau planaire: nombre chromatique majoré par 4: soit une espérance du nombre d'étapes: $e^\gamma 4 \simeq 7.12$

Conséquence, cas distribué

On pose d_i le degré du joueur i , i.e. le nombre de joueurs non indépendants avec i , et d^* le degré maximal.

- Collision entre deux joueurs indépendants sans effet.
- Probabilité joueur i cause une collision avec un voisin:
 $1 - \exp(-\frac{\lambda}{n} d_i \delta)$ au lieu de $1 - \exp(-\lambda \delta)$.
- On peut accélérer le taux de chaque joueur à $\frac{\hat{\lambda}}{d^*}$ au lieu de $\frac{\lambda}{n}$, ce qui donne un taux global de $\frac{n}{d^*} \lambda$ tout en réduisant les collisions.
- La complexité devient bornée par $\delta e^\gamma d^* \frac{\log n}{\log \log n}$.
- Dans le cas d'un réseau de degré fini, cela donne un temps d'exécution de $O(\log n)$

En récapitulant

- l'analyse de la Dynamique de Meilleure Réponse et variante
 - Cas séquentiel: Linéaire en moyenne contre exponentielle en pire cas
 - Cas distribué: Possible sans avoir à garantir un jeu séquentiel. complexité quasi linéaire ($O(n \log n)$) en moyenne. Proportion de collisions optimale tend vers 1 lorsque le nombre de joueurs grandit.
 - Cas structuré: gain important, potentiellement temps constant.
- Ces résultats sont obtenus par une approche analytique.

Extensions possibles

- L'approche markovienne reste valide pour d'autres critères (variance et autres moments, temps médian, mode, probabilité de dépasser un temps, borne à 95%)
- Améliorer les bornes naïves obtenues dans le cas en réseau et étudier d'autres types de structures.
- Application d'une approche similaire à d'autres problèmes de théorie des jeux
→ forts écarts pire cas - cas pratique dans de nombreux problèmes