Théorie des jeux - Dilemme itéré des prisonniers

Sébastien Konieczny

konieczny@irit.fr

IRIT-CNRS

Université Paul Sabatier - Toulouse

D’après le travail de Bruno Beaufils

beaufils@lifl.fr

http://www.lifl.fr/IPD

La petite histoire des prisonniers...

Deux personnes arrêtées ensemble en possession d’armes à feu sont soupçonnés d’un délit fait en commun. Les policiers les séparent et disent à chacun :

- Si un des deux avoue et que l’autre n’avoue rien, le premier est libéré, et le second emprisonné (5 ans);
- Si les deux avouent, les deux iront en prison (3 ans);
- Si aucun des deux n’avoue, les deux seront seront libérés assez vite (1 an).

    - Vous êtes un des deux prisonniers, que faites-vous ?

Le dilemme itéré...

Vous n'avez pas vraiment les mêmes goûts que votre voisin en matière de musique. Il lui arrive souvent d'écouter de la techno à fond. De même il vous arrive (en représailles) de mettre votre musique à un volume plus que raisonnable. Ce qui a pour conséquences que le lendemain il recommence à nouveau. En dehors de ces périodes agitées, vous appréciez les périodes où aucun de vous ne gêne l'autre.
Supposons que l'on pondère votre satisfaction :

- ▷ Vous avez une satisfaction de **5** à écouter votre musique à un volume important.
- ▷ La satisfaction est de **0** lorsque votre voisin met sa musique à fond.
- ▷ Une soirée "calme", sans musique vous apporte une satisfaction de **3**.
- ▷ Le fait d'écouter "simultanément" votre musique mêlée à celle du voisin, donne une satisfaction de **1**.

  - ▷ Vous savez ce que votre voisin à eu comme comportement les jours précédents, que faites-vous aujourd'hui?

Le Dilemme . . .

- Introduction par FLOOD et DRESHER à la RAND Corp. en 1952
- Jeu à somme non-nulle
- 2 joueurs jouent simultanément
- 2 choix de jeux :
  - COOPÉRER, *i.e.* être gentil, on notera `C`
  - TRAHIR, *i.e.* être méchant, on notera `D`
- Les gains des joueurs, notés S, P, R et T, sont fonction de leur choix de jeu avec :

$$S < P < R < T$$

Le Dilemme Itéré . . .

- ▷ Les joueurs se rencontrent plusieurs fois
- ▷ Ils ne connaissent pas le terme du jeu
- ▷ Le gain d’un joueur est le cumul de ses gains dans chaque rencontre
- ▷ Pour favoriser la coopération on rajoute la contrainte :

$$S + T < 2R$$

Dilemme Itéré des Prisonniers (résumé)

Dilemme… $S < P < R < T$

…itéré $S + T < 2R$

	Cooperate	Defect
Cooperate	$R = 3$ Reward récompense pour coopération mutuelle	$S = 0$ Sucker's payoff salaire de la dupe
Defect	$T = 5$ Temptation Tentation à trahir	$P = 1$ Punishment punition pour la trahison mutuelle

Score du joueur de la ligne.

Des applications concrètes...

- Deux pays doivent-ils lever des taxes douanières sur les produits importés de l’autre pays.
- Deux entreprises concurrentes doivent-elles essayer de s’entendre pour se partagé un marché ou se faire concurrence ?
- Deux espèces vivant sur un même territoire doivent-elles cohabiter ou se disputer la nourriture disponible ?

Les stratégies

- ▷ À chaque itération les joueurs ont connaissance des coups précédents
- ▷ Les joueurs ne peuvent pas passer d'accord
- ▷ Le nombre de parties n'est pas connu à l'avance

Quelques exemples :

- ▷ gentille
- ▷ méchante
- ▷ per_CCD
- ▷ rancunière
- ▷ lunatique
- ▷ majoritaire_gentille
- ▷ donnant_donnant
- ▷ graduelle

Exemples (rencontres)

	1	2	3	4	5	6	7	8	9	10		
score de gentille	0	0	0	0	0	0	0	0	0	0	=	0
jeu de gentille	C	C	C	C	C	C	C	C	C	C		
jeu de méchante	D	D	D	D	D	D	D	D	D	D		
score de méchante	5	5	5	5	5	5	5	5	5	5	=	50

	1	2	3	4	5	6	7	8	9	10		
score de per_CCD	3	3	5	0	0	1	0	0	1	0	=	13
jeu de per_CCD	C	C	D	C	C	D	C	C	D	C		
jeu de rancunière	C	C	C	D	D	D	D	D	D	D		
score de rancunière	3	3	0	5	5	1	5	5	1	5	=	33

Les tournois

- Plusieurs stratégies se rencontrent 2 à 2, comme pour un tournoi sportif
- Le gain d’une stratégie est le cumul de ses scores face à chaque adversaire
- Toutes les parties ont la même longueur (même nombre d’itérations), mais les stratégies ne la connaissent pas et ne peuvent pas le savoir

Exemples (tournoi)

	gentille	méchante	per_CCD	rancunière
gentille	30	50	36	30
méchante	0	10	3	9
per_CCD	21	38	24	33
rancunière	30	14	13	30

Score	81	112	76	102

Classement
1. méchante
2. rancunière
3. gentille
4. per_CCD

Quelle est la *meilleure* stratégie ?

- qui batte toutes les autres :

Quelle est la *meilleure* stratégie ?

- qui batte toutes les autres :
  méchante, car généralisation du dilemme non itéré

Quelle est la *meilleure* stratégie ?

- qui batte toutes les autres :
  méchante, car généralisation du dilemme non itéré
- qui fasse le meilleur score possible face à toutes les autres :

Quelle est la *meilleure* stratégie ?

- qui batte toutes les autres :
  méchante, car généralisation du dilemme non itéré
- qui fasse le meilleur score possible face à toutes les autres :
  aucune, car meilleure contre méchante et contre rancunière est impossible

Quelle est la *meilleure* stratégie ?

- qui batte toutes les autres :
  méchante, car généralisation du dilemme non itéré
- qui fasse le meilleur score possible face à toutes les autres :
  aucune, car meilleure contre méchante et contre rancunière est impossible

  - Problème de définition du critère d'évaluation des stratégies

Quelle est la *meilleure* stratégie ?

Sur des confrontations de 100 parties :

- ▷ Le gain maximal est de 500 points
- ▷ Le gain minimal est de 0 point

C'est ce qu'obtiennent MÉCHANTE et GENTILLE l'une contre l'autre.

Quelle est la *meilleure* stratégie ?

Sur des confrontations de 100 parties :

- ▷ Le gain maximal est de 500 points
- ▷ Le gain minimal est de 0 point

C'est ce qu'obtiennent MÉCHANTE et GENTILLE l'une contre l'autre.
Mais...

- ▷ 2 gentilles entre elles obtiennent chacune 300 points
- ▷ 2 méchantes entre elles obtiennent chacune 100 points
- ▷ Chaque stratégie est bonne (au sens du meilleur score) face à certaines et mauvaises face à d'autres car elle ne sait pas à qui elle a affaire.

donnant-donnant : une bonne stratégie

Au premier coup je coopère (`C`), ensuite si mon adversaire a coopéré (`C`) au coup précédent, je coopère (`C`), s'il a trahi (`D`), je trahis (`D`).

donnant-donnant : une bonne stratégie

Au premier coup je coopère (`C`), ensuite si mon adversaire a coopéré (`C`) au coup précédent, je coopère (`C`), s'il a trahi (`D`), je trahis (`D`).

- ▷ donnant-donnant ne gagne jamais contre personne !

donnant-donnant : une bonne stratégie

Au premier coup je coopère (`C`), ensuite si mon adversaire a coopéré (`C`) au coup précédent, je coopère (`C`), s'il a trahi (`D`), je trahis (`D`).

- ▷ donnant-donnant ne gagne jamais contre personne !
- ▷ Au mieux elle fait le même score.

donnant-donnant : une bonne stratégie

Au premier coup je coopère (`C`), ensuite si mon adversaire a coopéré (`C`) au coup précédent, je coopère (`C`), s'il a trahi (`D`), je trahis (`D`).

- donnant-donnant ne gagne jamais contre personne !
- Au mieux elle fait le même score.
- Mais, au pire elle ne perd que 5 points quel que soit l'adversaire et la longueur de la partie !

Un tournoi

Tournois entre 10 stratégies parmi 12 :

- gentille
- méchante
- lunatique
- donnant_donnant
- rancunière
- per_DDC
- per_CCD
- majoritaire_gentille
- majoritaire_méchante
- méfi ante
- sondeur
- donnant_donnant_dur

Nombre de tournois joués par chaque stratégie : 55

- Donnez le classement du tournoi...

Un tournoi

- gentille
- méchante
- lunatique
- donnant_donnant
- rancunière
- per_DDC
- per_CCD
- majoritaire_gentille
- majoritaire_méchante
- méfi ante
- sondeur
- donnant_donnant_dur

Scores :

donnant_donnant	:	40
majoritaire_gentille	:	17
rancunière	:	7
majoritaire_méchante	:	2
lunatique	:	0
méchante	:	0

Évolution écologique

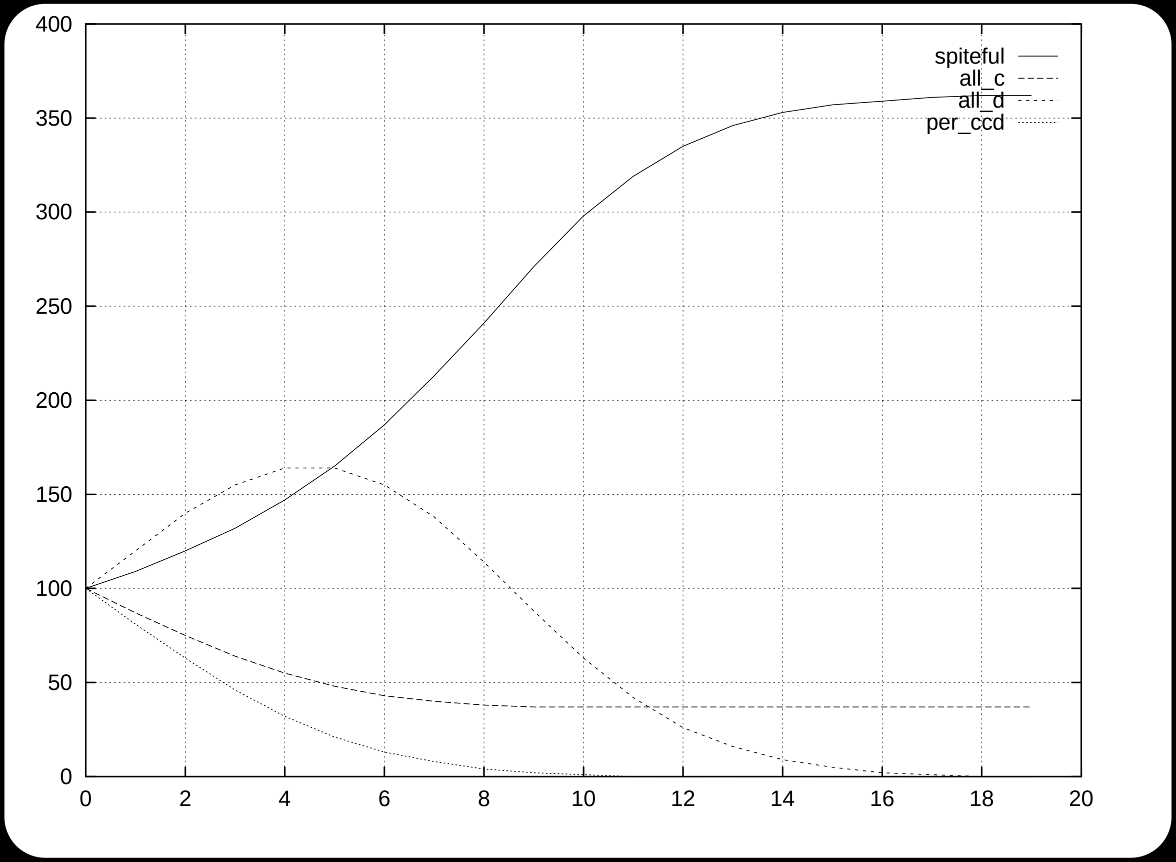
Simulation de l'évolution naturelle :

- ▷ Chaque stratégie est représentée par une population de N entités
- ▷ Un tournoi entre toutes les entités est fait
- ▷ Les entités de faibles stratégies (au sens du classement dans le tournoi) sont défavorisées, celles à stratégie forte sont favorisées
- ▷ La favorisation est faite par une redistribution proportionnelle de la population

Ce cycle est répété jusqu'à stabilisation de la population

Exemples (évolution)

	1	2	3	4	...
gentille	100	81	87	37	...
méchante	100	112	120	42	...
rancunière	100	102	109	319	...
per_CCD	100	76	81	0	...



Une morale très morale...

Critères de qualité pour une stratégie (en évolution) : [Axelrod,81]

- *Gentillesse*
- *Réactivité*
- *Pardon*
- *Simplicité*

Une morale très morale...

Critères de qualité pour une stratégie (en évolution) : [Axelrod,81]

- ▷ *Gentillesse*
- ▷ *Réactivité*
- ▷ *Pardon*
- ▷ *Simplicité*

Les bonnes stratégies au dilemme le sont aussi dans les variantes du dilemme (asynchrone, avec renoncement, bruits, . . .)
Mais la simplicité n'est peut-être pas un bon critère (exemple: graduelle) [Beaufils *et al*,96]

Une morale très morale...

Critères de qualité pour une stratégie (en évolution) : [Axelrod,81]

- ▷ *Gentillesse*
- ▷ *Réactivité*
- ▷ *Pardon*
- ▷ *Simplicité*

Les bonnes stratégies au dilemme le sont aussi dans les variantes du dilemme (asynchrone, avec renoncement, bruits, . . .)
Mais la simplicité n'est peut-être pas un bon critère (exemple: graduelle) [Beaufils *et al*,96]
Deux gros problèmes :

- ▷ Comment automatiser la recherche d'une stratégie
- ▷ Comment évaluer le plus objectivement possible une stratégie

Définir une classe de stratégies

- Pour automatiser la recherche il faut déterminer une méthode descriptive de définir un ensemble de stratégies
- Pour définir un ensemble de stratégies on peut par exemple :
  - définir une structure capable d'être décodée en un comportement à adopter face à un adversaire
  - utiliser toutes les manières possibles de remplir cette structure comme autant de stratégies

$\Longrightarrow$ approche génétique de définir des individus

Méthodes exhaustives risquent :

- de ne pas être objective
- jamais complètes
- inutilisable en retour (incompréhension des traits de la stratégie)

Définir une classe de stratégies

Considérons les stratégies qui ne se souviennent que de leur dernier coup et du dernier coup de leur adversaire.

On peut définir une de celles-ci par :

- le premier coup je joue $\boxed{\texttt{C}}$ puis
- $\begin{cases} \text{si au coup d'avant j'ai joué } \texttt{C} \text{ et que lui a joué } \texttt{C} \text{ alors je joue } \boxed{\texttt{C}} \\ \text{si au coup d'avant j'ai joué } \texttt{C} \text{ et que lui a joué } \texttt{D} \text{ alors je joue } \boxed{\texttt{D}} \\ \text{si au coup d'avant j'ai joué } \texttt{D} \text{ et que lui a joué } \texttt{C} \text{ alors je joue } \boxed{\texttt{C}} \\ \text{si au coup d'avant j'ai joué } \texttt{D} \text{ et que lui a joué } \texttt{D} \text{ alors je joue } \boxed{\texttt{D}} \end{cases}$

Définir une classe de stratégies

Considérons les stratégies qui ne se souviennent que de leur dernier coup et du dernier coup de leur adversaire.

On peut définir une de celles-ci par :

- le premier coup je joue $\boxed{\texttt{C}}$ puis
- $$\begin{cases} \text{si au coup d'avant j'ai joué } \texttt{C} \text{ et que lui a joué } \texttt{C} \text{ alors je joue } \boxed{\texttt{C}} \\ \text{si au coup d'avant j'ai joué } \texttt{C} \text{ et que lui a joué } \texttt{D} \text{ alors je joue } \boxed{\texttt{D}} \\ \text{si au coup d'avant j'ai joué } \texttt{D} \text{ et que lui a joué } \texttt{C} \text{ alors je joue } \boxed{\texttt{C}} \\ \text{si au coup d'avant j'ai joué } \texttt{D} \text{ et que lui a joué } \texttt{D} \text{ alors je joue } \boxed{\texttt{D}} \end{cases}$$

On peut alors définir le génotype de cette stratégie comme :

C	C	D	C	D

Définir une classe de stratégies

Considérons les stratégies qui ne se souviennent que de leur dernier coup et du dernier coup de leur adversaire.

On peut définir une de celles-ci par :

- ▷ le premier coup je joue $\boxed{\texttt{C}}$ puis
- ▷ $\left\{ \begin{array}{l} \text{si au coup d'avant j'ai joué } \texttt{C} \text{ et que lui a joué } \texttt{C} \text{ alors je joue } \boxed{\texttt{C}} \\ \text{si au coup d'avant j'ai joué } \texttt{C} \text{ et que lui a joué } \texttt{D} \text{ alors je joue } \boxed{\texttt{D}} \\ \text{si au coup d'avant j'ai joué } \texttt{D} \text{ et que lui a joué } \texttt{C} \text{ alors je joue } \boxed{\texttt{C}} \\ \text{si au coup d'avant j'ai joué } \texttt{D} \text{ et que lui a joué } \texttt{D} \text{ alors je joue } \boxed{\texttt{D}} \end{array} \right.$

On peut alors définir le génotype de cette stratégie comme :

C	C	D	C	D

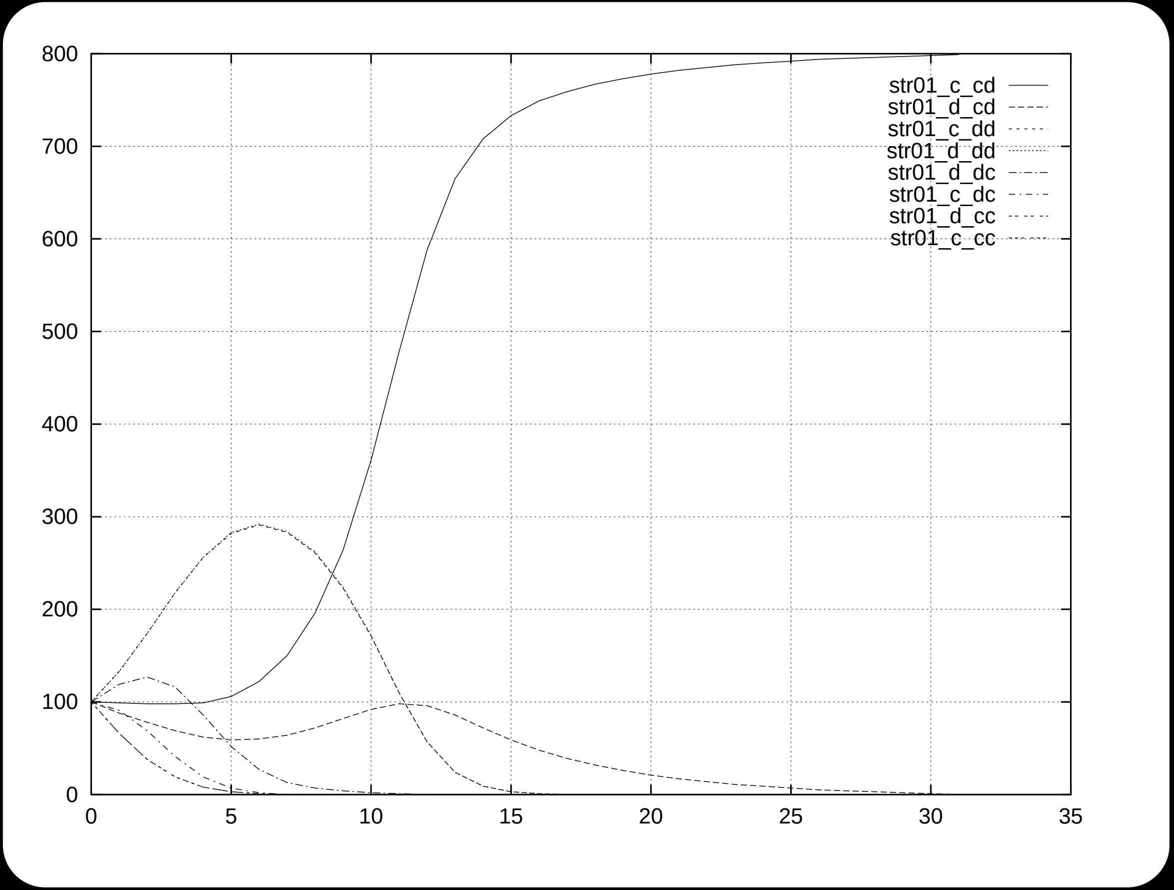
▷ C'est donnant_donnant

La classe memory

- Chaque stratégie ne voit que M coups de son passé, et O coups du passé de son adversaire
- La stratégie amorce son jeu par $\max(M, O)$ coups prédéfinis
- La réponse à un coup de l'adversaire ne dépend que de la configuration historique visible
- Le génotype a donc une longueur de $\max(M, O) + 2^{(M+O)}$

La classe memory (exemple)

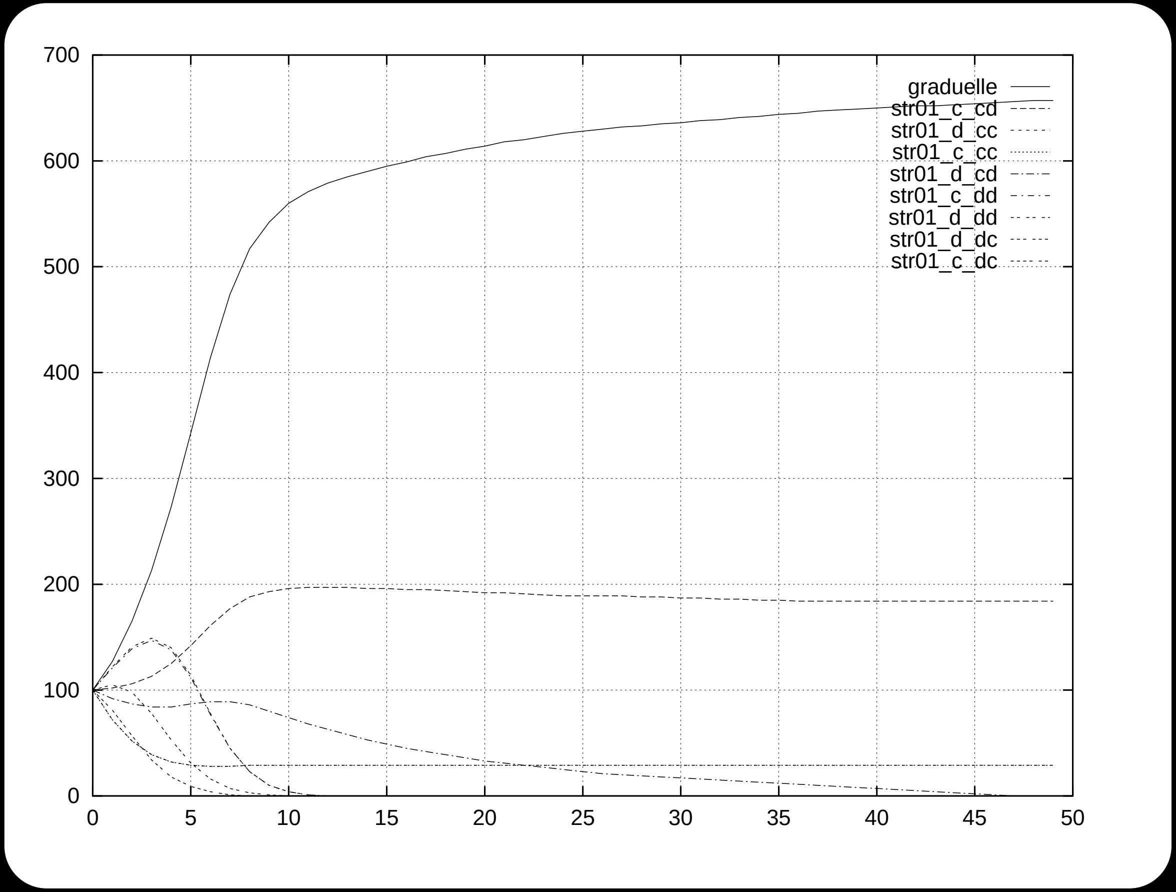
$M = 0$ et $O = 1$, soit 8 stratégies



Le vainqueur est donnant_donnant

La classe memory + graduelle

$M = 0$, $O = 1$ et graduelle est dans la population



Quelques résultats

Pour évaluer une stratégie on fait des évolutions de classes complètes en ajoutant celle-ci dans la population de départ, puis on regarde son classement à la fin de l'évolution.

Par exemple évaluons graduelle, donnant_donnant et rancunière:

	M	O	taille	graduelle	donnant_don.	rancunière
memory	0	1	8	1	1	1
	0	2	64	5	2	21
	1	1	32	2	3	1
	1	2	1024	6	13	37
binary_memory	0	1	32	1	2	1
	1	1	512	1	7	13
memory_automata	0	1	512	1	31	13

Tout sur le dilemme...

Pour plus de détails sur le dilemme itéré des prisonniers :

`http://www.lifl.fr/IPD`

- ▷ Bruno BEAUFILS
- ▷ Jean-Paul DELAHAYE
- ▷ Philippe MATHIEU

Laboratoire d'Informatique Fondamentale de Lille

Rome-Alésia

Rome						Alésia
□	○	○	●	○	○	□
50						50

Les règles :

- ▷ Deux villes sont en guerre
- ▷ Le champ de bataille initial est représenté par le ●
- ▷ Chaque ville à initialement 50 armées
- ▷ Pour chaque bataille, chaque ville envoie (dépense) un certain nombre d’armées (toutes les armées dépensées sont perdues)
- ▷ La ville qui envoie le plus grand nombre d’armées gagne la bataille, le champ de bataille se déplace vers la ville adverse. Si les deux villes envoie un même nombre d’armées le champ de bataille ne bouge pas
- ▷ Le vainqueur est la ville qui réussit à envahir la ville adverse

exemple

Rome						Alésia
□	○	○	●	○	○	□
50						50

exemple

Rome						Alésia
□	○	○	●	○	○	□
50						50

Rome 5 - 10 Alésia

exemple

Rome Alésia

□ ○ ○ ● ○ ○ □

50 50

Rome 5 - 10 Alésia

Rome Alésia

□ ○ ● ○ ○ ○ □

45 40

exemple

Rome Alésia

□ ○ ○ ● ○ ○ □

50 50

Rome 5 - 10 Alésia

Rome Alésia

□ ○ ● ○ ○ ○ □

45 40

Rome 10 - 15 Alésia

Rome Alésia

□ ● ○ ○ ○ ○ □

35 25

exemple

Rome						Alésia
□	○	○	●	○	○	□
50						50

Rome 5 - 10 Alésia

Rome						Alésia
□	○	●	○	○	○	□
45						40

Rome 10 - 15 Alésia

Rome						Alésia
□	●	○	○	○	○	□
35						25

Rome 25 - 1 Alésia

Rome						Alésia
□	○	●	○	○	○	□
10						24

exemple

Rome Alésia

□ ○ ○ ● ○ ○ □

50 50

Rome 5 - 10 Alésia

Rome Alésia

□ ○ ● ○ ○ ○ □

45 40

Rome 10 - 15 Alésia

Rome Alésia

□ ● ○ ○ ○ ○ □

35 25

Rome 25 - 1 Alésia

Rome Alésia

□ ○ ● ○ ○ ○ □

10 24

Rome 2 - 11 Alésia

Rome Alésia

□ ● ○ ○ ○ ○ □

8 13

exemple

Rome Alésia
□ ○ ○ ● ○ ○ □
50 50
Rome 5 - 10 Alésia

Rome Alésia
□ ○ ● ○ ○ ○ □
45 40
Rome 10 - 15 Alésia

Rome Alésia
□ ● ○ ○ ○ ○ □
35 25
Rome 25 - 1 Alésia

Rome Alésia
□ ○ ● ○ ○ ○ □
10 24
Rome 2 - 11 Alésia

Rome Alésia
□ ● ○ ○ ○ ○ □
8 13
Rome 4 - 9 Alésia

Rome Alésia
■ ○ ○ ○ ○ ○ □
4 4
Vainqueur : Alésia !

Itération

But : Comprendre le comportement de l’adversaire.

Itération

But : Comprendre le comportement de l'adversaire.

Quelques stratégies possibles:

- ▷ **Statistique :** au $n$ieme coup mon adversaire a joué en moyenne k unités, je vais jouer $k+1$.

Itération

But : Comprendre le comportement de l'adversaire.

Quelques stratégies possibles:

- **Statistique :** au $n$ieme coup mon adversaire a joué en moyenne k unités, je vais jouer $k+1$.
- **Règles :** Si mon adversaire est à une case de la victoire alors jouer k unités, où k est le nombre d'unités qu'il lui reste.

Itération

But : Comprendre le comportement de l'adversaire.

Quelques stratégies possibles:

- **Statistique :** au $n$ieme coup mon adversaire a joué en moyenne k unités, je vais jouer $k+1$.
- **Règles :** Si mon adversaire est à une case de la victoire alors jouer k unités, où k est le nombre d'unités qu'il lui reste.
- **Probabiliste :** Au premier coup
  jouer 1 avec une probabilité de $0,1$
  jouer 7 avec une probabilité de $0,4$
  jouer 12 avec une probabilité de $0,3$
  le reste du temps jouer au hasard **(aléatoire)**

Itération

But : Comprendre le comportement de l'adversaire.

Quelques stratégies possibles:

- ▷ **Statistique :** au $n$ieme coup mon adversaire a joué en moyenne k unités, je vais jouer $k+1$.
- ▷ **Règles :** Si mon adversaire est à une case de la victoire alors jouer k unités, où k est le nombre d'unités qu'il lui reste.
- ▷ **Probabiliste :** Au premier coup
  jouer 1 avec une probabilité de $0,1$
  jouer 7 avec une probabilité de $0,4$
  jouer 12 avec une probabilité de $0,3$
  le reste du temps jouer au hasard **(aléatoire)**

Une bonne stratégie doit :

- ▷ **Comprendre** le comportement de son adversaire
- ▷ **Cacher** son propre comportement