Théorie des graphes, et de l'approximation pour le routage, la coloration et l'apprentissage d'équilibres.

Johanne Cohen

PRiSM/CNRS

Travail commun avec O. Bournez (LIX, école polytechnique)

Plan de présentation

1. Définition des jeux de ressources
   1.1 Un jeu d'ordonnancement
   1.2 Un jeu de congestion
2. Apprentissage des équilibres de Nash.
   2.1 La dynamique de la meilleure réponse
   2.2 Une classe d'algorithmes d'apprentissage.
3. Apprentissage basé sur la dynamique de réplication
4. Conclusion

1. Jeux de ressources

1.1. Jeu d'ordonnancement

1.2. Jeu de congestion

Jeux de partage de ressources I

- n joueurs en compétition sur un ensemble de m ressources.
- chaque ressource possède une fonction de coût croissante :

$$C_j : \mathbb{N} \rightarrow \mathbb{R},\ 1 \leq j \leq m$$

Jeux d'ordonnancement

[Koutsoupias, Papadimitriou 99]

- chaque joueur i = une tâche de longueur w_i
- chaque ressource j = une machine avec une vitesse s_j
- coût du joueur i = la charge de la machine m_i hébergeant i

soit $c_{m_i}\left(\frac{\sum_{j|m_k=m_i} w_k}{s_{m_i}}\right)$

Jeux de partage de ressources I

- n joueurs en compétition sur un ensemble de m ressources.
- chaque ressource possède une fonction de coût croissante :

$$C_j : \mathbb{N} \to \mathbb{R},\ 1 \leq j \leq m$$

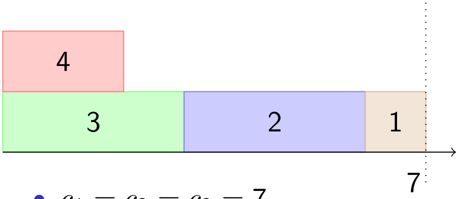
Jeux d'ordonnancement

[Koutsoupias, Papadimitriou 99]

- chaque joueur i = une tâche de longueur w_i
- chaque ressource j = une machine avec une vitesse s_j
- coût du joueur i = la charge de la machine m_i hébergant i

soit $c_{m_i}(\frac{\sum_{j|m_k=m_i} w_k}{s_{m_i}})$

4 tâches ($w_1 = 1$, $w_2 = w_3 = 3$, $w_4 = 2$)
2 machines identiques



- $c_1 = c_2 = c_3 = 7$
- $c_4 = 2$

Jeux de partage de ressources I

- n joueurs en compétition sur un ensemble de m ressources.
- chaque ressource possède une fonction de coût croissante :

$$C_j : \mathbb{N} \to \mathbb{R},\ 1 \le j \le m$$

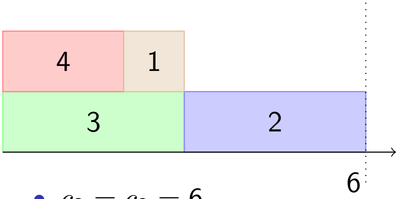
Jeux d'ordonnancement

[Koutsoupias, Papadimitriou 99]

4 tâches ($w_1 = 1$, $w_2 = w_3 = 3$, $w_4 = 2$)
2 machines identiques

- chaque joueur i = une tâche de longueur w_i
- chaque ressource j = une machine avec une vitesse s_j
- coût du joueur i = la charge de la machine m_i hébergant i

soit $c_{m_i}\left(\frac{\sum_{j|m_k=m_i} w_k}{s_{m_i}}\right)$

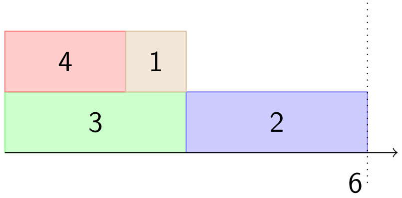


- $c_2 = c_3 = 6$
- $c_4 = c_1 = 3$

Un équilibre de Nash

Un placement de tâches A est un équilibre de Nash si et seulement si

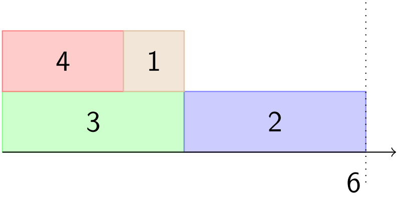
aucun joueur ne peut améliorer le coût unilatéralement en déplacement une tâche vers une autre machine



Un équilibre de Nash

Un placement de tâches A est un équilibre de Nash si et seulement si

Pour chaque joueur i, et pour toutes les machines k, nous avons : $c_i^{A[i]} \leq c_i^k$



Jeux de partage de ressources (II)

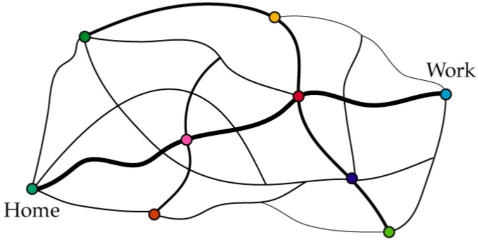
- n joueurs en concurrence sur un ensemble de m ressources.
- chaque ressource possède une fonction de charge croissante

$$C_j : \mathbb{N} \to \mathbb{R},\ 1 \leq j \leq m$$

Jeux de congestion

[Beckmann, McGuire, Winsten 56]

- chaque joueur i = un couple de sommets (H_i, W_i) d'un graphe G.
- chaque ressource j = un arête du graphe
- coût du joueur i = la somme des charges des ressources du chemin choisi entre H_i et W_i



▸ détail

Les jeux de potentiel

[Rosenthal 73]

Un jeu est dit un jeu de potentiel s'il existe une fonction ϕ dont les variations reflètent les variations individuelles de coût.

Formellement : pour tout joueur i, pour toute paire de stratégies pures k et k' de i

pour le cas exact :

$$\phi(Q_{-i}, k') - \phi(Q_{-i}, k) = c_i(Q_{-i}, k) - c_i(Q_{-i}, k').$$

pour le cas ordinal :

$$signe(\phi(Q_{-i}, k') - \phi(Q_{-i}, k)) = signe(c_i(Q_{-i}, k) - c_i(Q_{-i}, k'))$$

Résultats

1. [Rosenthal 73] Tout jeu de potentiel ordinal possède toujours un équilibre de Nash pur.

2. Faits connus :

    - Les jeux de congestion, et les jeux d'ordonnancement sont des jeux de potentiel exact.

    - Un jeu de potentiel exact est un jeu de potentiel ordinal.

3. [Monderer et al 96] Tout jeu de potentiel exact est isomorphe à un jeu de congestion.

2. Apprentissage des équilibres de Nash

2.1. La dynamique de la meilleure réponse

2.2. Une classe d'algorithmes d'apprentissage

Apprentissage d'un équilibre de Nash.

- Un ensemble de joueurs : $1 \leq i \leq n$

- Chaque joueur i a un ensemble fini de m_i stratégies pures.

- A chaque étape t , le joueur i choisit une stratégie pure selon le vecteur de probabilité $q_i(t)$

$$q_{i,\ell}(t) \in [0, \dots, 1], \quad \sum_{\ell=1}^{m_i} q_{i,\ell}(t) = 1.$$

- Algorithme d'apprentissage :

$$q_i(t+1) = Algorithme(\{q_j(t')\}_{j \text{ joueur}, t' \leq t})$$

tel que $Q(t) = (q_1(t), \cdots, q_n)$ converge vers un équilibre de Nash, lorsque $t \to \infty$.

Dynamiques de meilleures réponses

- chaque joueur qui peut améliorer son coût joue l'un après l'autre ;
- Lorsque c'est son tour, chaque joueur i change de stratégie pour une qui minimise son coût.

(Version asynchrone)

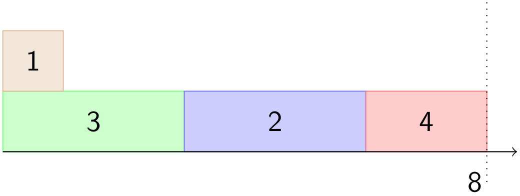
- [Monderer et al 96] Toute dynamique de meilleure réponse atteint ultimement un équilibre de Nash pur pour les jeux de potentiel.

- Bornes sur le temps de convergence :
  - [Even-Dar et al 07] pour les jeux d'ordonnancement
  - [Chien et al 09] pour les jeux de congestion.

Dynamiques de meilleures réponses : Exemple

Theorem [Even-Dar2003]

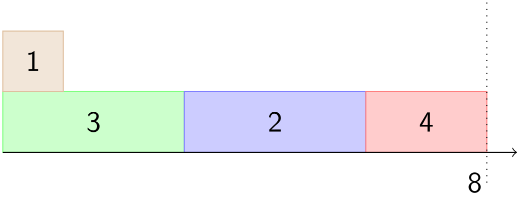
Soit A un placement de n tâches sur m machines identiques. Commençant de A, la dynamique de meilleure réponse atteint ultimement un équilibre de Nash pur.



Dynamiques de meilleures réponses : Exemple

Theorem [Even-Dar2003]

Soit A un placement de n tâches sur m machines identiques. Commençant de A, la dynamique de meilleure réponse atteint ultimement un équilibre de Nash pur.

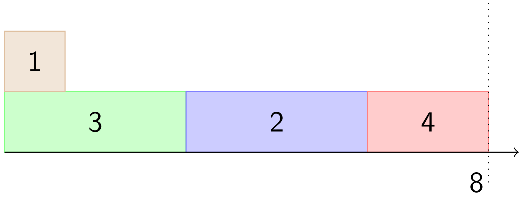


Les joueurs 2, 3 et 4 peuvent améliorer leur coût

Dynamiques de meilleures réponses : Exemple

Theorem [Even-Dar2003]

Soit A un placement de n tâches sur m machines identiques. Commençant de A, la dynamique de meilleure réponse atteint ultimement un équilibre de Nash pur.

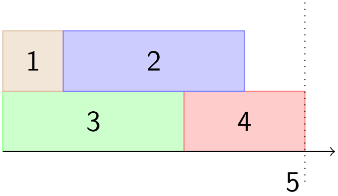


C'est le tour du joueur 2

Dynamiques de meilleures réponses : Exemple

Theorem [Even-Dar2003]

Soit A un placement de n tâches sur m machines identiques. Commençant de A, la dynamique de meilleure réponse atteint ultimement un équilibre de Nash pur.



Placement de tâches après le déplacement

Dynamique de la meilleure réponse seuillée.

Version où les joueurs sont identiques

1. A l'initialisation,

   chaque joueur i a une stratégie pure.

2. A chaque étape t, pour chaque joueur i, faire en parallèle,
   - Soit r_i la ressource courante de la tâche i.
   - Choisir une autre ressource de façon uniforme.
   - Si $C_{r_i}(t) \geq C_j(t)$ alors déplacer la tâche i vers la ressource j avec la probabilité $(1 - \frac{C_j(t)}{C_{r_i}(t)})$.

Theorem [Berenbrink2006]

Cette dynamique atteint ultimmement un équilibre de Nash pur en un nombre d'étapes espéré $O(\log \log n + m^4)$ lorsque toutes les joueurs sont identiques.

2. Apprentissage des équilibres de Nash

2.1. La dynamique de la meilleure réponse

2.2. Une classe d'algorithmes d'apprentissage

Une classe d'algorithmes d'apprentissage.

1. A l'initialisation (à l'étape 0),

   $q_i(0)$ est une stratégie mixte pour tout joueur i.

2. Pour chaque étape t,

   2.1 Chaque joueur i sélectionne une stratégie $s_i(t)$ selon $\sigma_i(q_i(t))$.

   2.2 Un coût $c_i(t)$ pour chaque joueur i est obtenu par ces choix

   2.3 Chaque joueur i

   réactualise $q_i(t)$:

$$q_i(t+1) = q_i(t) + \mathrm{b}F_i^b(c_i(t), s_i(t), q_i(t)).$$

Version synchrone / Version asynchrone

F_i^b : fonction continue, conservation de vecteur de probabilité.

$G_i(Q) = \lim_{b \to 0} \mathbb{E}[\ F_i^b(c_i(t), s_i(t), q_i(t))\ |\sigma(Q(t))\ ]$ existe

Une classe d'algorithmes d'apprentissage.

1. A l'initialisation (à l'étape 0),

   $q_i(0)$ est une stratégie mixte pour tout joueur i.

2. Pour chaque étape t,

   2.1 Chaque joueur i sélectionne une stratégie $s_i(t)$ selon $\sigma_i(q_i(t))$.

   2.2 Un coût $c_i(t)$ pour chaque joueur i est obtenu par ces choix

   2.3 Un seul joueur i est choisit aléatoirement, et il

   réactualise $q_i(t)$:

$$q_i(t+1) = q_i(t) + \mathbf{b} F_i^b(c_i(t), s_i(t), q_i(t)).$$

Version synchrone /Version asynchrone

F_i^b : fonction continue, conservation de vecteur de probabilité.

$G_i(Q) = \lim_{b \to 0} \mathbb{E}[\ F_i^b(c_i(t), s_i(t), q_i(t))\ |\sigma(Q(t))\ ]$ existe

Méthode de l'EDO I

L'algorithme est une discrétisation (stochastique)
d'une équation différentielle ordinaire

- Les trajectoires de l'algorithme stochastique convergent faiblement lorsque $b \to 0$ vers la solution de l'équation différentielle ordinaire

$$\left\{ \begin{array}{lcl} \frac{dX}{dt} & = & G(X) \\ X(0) & = & Q(0) \end{array} \right.$$

Une suite Q_n de mesures converge faiblement vers Q, si $E[f(Q_n)] \to E[f(Q)]$ pour toute fonction $f(.)$ continue et bornée.

- En considérant

1. σ_i comme les fonctions identités ;
   $F_i^b(c_i(t), s_i(t), q_i(t)) = c_i(t)(e_{s_i(t)} - q_i(t))$

   on obtient une dynamique de réplication.

2. $\sigma_{i\ell}(q_i) = \frac{\exp(q_{i,\ell}/\kappa)}{\sum_j \exp(q_{i,j}/\kappa)}$
   $F_i^b(c_i(t), s_i(t), q_i(t)) = (c_i(t) - q_{i,s_i(t)})e_{s_i(t)}$

   on obtient une dynamique de *logit*.

3. ...

   on obtient une dynamique de la meilleure réponse seuillée.

Méthode de l'EDO II

Theorem

Borner le temps de convergence de l'algorithme stochastique correspond exactement à borner le temps de convergence de l'équation différentielle ordinaire associée.

3. Apprentissage basé sur la dynamique de réplication

La dynamique de réplication pertubée.

1. A l'initialisation (à l'étape 0),

   $q_i(0)$ est une stratégie mixte pour tout joueur i.

2. Pour chaque étape t,

   2.1 Chaque joueur i sélectionne une stratégie $s_i(t)$ selon $q_i(t)$.

   2.2 Un coût $c_i(t)$ pour chaque joueur i est obtenu par ces choix

   2.3 Un seul joueur i est choisit aléatoirement, et il

   réactualise $q_i(t)$:

$$q_i(t+1) = q_i(t) + \mathbf{b} c_i(t) \times (\mathbf{e}_{a_i} - q_i(t)) + \mathcal{O}(b^2)$$

avec $\mathbf{e_{i\ell}} = \begin{cases} 1 & si\ \ell = i \\ 0 & sinon \end{cases}$

Pourquoi cet algorithme.

- Soit $Q(t) = (q_1, q_2, \cdots, q_n)(t)$.
- Notons

$$h_{i,\ell}(Q) = \mathbb{E}[c_i | i \text{ joue } \ell, \text{ les autres jouent selon } Q]$$

$$\overline{h_i}(Q) = \mathbb{E}[c_i | \text{ tous les joueurs jouent selon } Q]$$

Theorem

Les trajectoires de l'algorithme précédent convergent faiblement vers les solutions de

$$\frac{dq_{i,\ell}}{dt} = -q_{i,\ell}[h_{i,\ell}(Q) - \overline{h_i}(Q)], \tag{1}$$

quand $b \to 0$.

Résultats formels

Theorem (Folk Theorem of Evolutionary Game Dynamics)

For ODE (1)

- *Corners are rest points of dynamics.*
- *Nash Equilibria are rest points.*
- *All rest points which are non-Nash Equilibrium are not stable.*

Donc **si** l'équation converge, alors elle converge vers un équilibre de Nash.

Cependant,

- Il n'existe aucune raison pour qu'elle converge dans le cas général.
- Si c'est le cas, il y a une double limite : $b \to 0$, $t \to \infty$.
- Vitesse de convergence ?

Résultat

Theorem

Dans un jeu de congestion, il y a presque surement convergence vers un équilibre de Nash de l'algorithme stochastique asynchrone.

discussion uniquement sur les jeux d'ordonnancement

Résultat

Theorem

Dans un jeu de congestion, il y a presque surement convergence vers un équilibre de Nash de l'algorithme stochastique asynchrone.

discussion uniquement sur les jeux d'ordonnancement

- La dynamique de Réplication

$$\frac{dq_{i,\ell}}{dt} = -q_{i,\ell}[h_{i,\ell}(Q) - \overline{h_i}(Q)], \qquad 1 \leq \ell \leq m_i, 1 \leq i \leq n$$

- converge pour ces types de classes
  - ils ont une fonction de potentiel F (fonction de Lyapunov) telle que $F(Q)$ décroit quand Q évolue [Sandholm 2001].
  - $$\frac{dF(Q(t))}{dt} \leq 0$$

Résultat

Theorem

Dans un jeu de congestion, il y a presque surement convergence vers un équilibre de Nash de l'algorithme stochastique asynchrone.

discussion uniquement sur les jeux d'ordonnancement

- La dynamique de Réplication

$$\frac{dq_{i,\ell}}{dt} = -q_{i,\ell}[h_{i,\ell}(Q) - \overline{h_i}(Q)], \qquad 1 \leq \ell \leq m_i, 1 \leq i \leq n$$

- converge pour ces types de classes
  - ils ont une fonction de potentiel F (fonction de Lyapunov) telle que $F(Q)$ décroit quand Q évolue [Sandholm 2001].
  - $$\frac{dF(Q(t))}{dt} \leq 0$$

Résultat

Theorem

Dans un jeu de congestion, il y a presque surement convergence vers un équilibre de Nash de l'algorithme stochastique asynchrone.

discussion uniquement sur les jeux d'ordonnancement

- La dynamique de Réplication

$$\frac{dq_{i,\ell}}{dt} = -q_{i,\ell}[h_{i,\ell}(Q) - \overline{h_i}(Q)], \qquad 1 \leq \ell \leq m_i, 1 \leq i \leq n$$

- converge pour ces types de classes
  - ils ont une fonction de potentiel F (fonction de Lyapunov) telle que $F(Q)$ décroit quand Q évolue [Sandholm 2001].
  - $$\frac{dF(Q(t))}{dt} \leq 0$$

Idée principale

- F : une fonction sur l'état global Q du système où

$$F(Q) = \sum_{k=1}^{m} 1/s_k \left[\frac{1}{2} (\sum_{j=1}^{N} q_{j,k} w_j)^2 + \sum_{j=1}^{N} q_{j,k} w_j^2 (1 - \frac{q_{j,k}}{2}) \right]$$

Propriété 1 : Chaque étape de cet algorithme change Q (un seul élément q_i) de la façon suivante

$$\Delta F(Q) = \mathbb{E}[F(Q(t+1)) - F(Q(t)|Q(t)]) \leq 0$$

Corollaire : F va décroitre en moyenne et elle va converger presque surement. Un point limite doit annuler ΔF.

Propriété 2 : Les équilibres de Nash purs sont des points limites (stable) ▸ Skip

Propriété 1 : $\Delta F \leq 0$

$$\Delta F = \mathbb{E}[F(Q(t+1)) - F(Q(t))|Q(t)]$$

satisfait

$$\Delta F = -\frac{b}{n}\sum_{i=1}^{n} w_i \sum_{k=1}^{m}\sum_{k'>k} q_{i,k}q_{i,k'}[h_{i,k} - h_{i,k'}]^2 + \mathcal{O}(b^2)$$

$$\Delta F \leq 0$$

quand b est petit.

Corollaire : F va décroire

Puisque $\Delta F = \mathbb{E}[F(Q(t+1)) - F(Q(t))|Q(t)] \leq 0$,

F va décroître presque surement.

Le théorème de Foster (Théorie des martingales) :
Soit $Q_0, Q_1, \cdots, Q_n, \cdots$ une suite de variables aléatoires ayant des valeurs dans E tel que pour certaine fonction $F : E \rightarrow \mathbb{R}^{\geq 0}$, $\mathbb{E}[F(Q_{n+1})|Q_n] \leq F(Q_n) - \epsilon$ chaque fois que $Q_n \notin S$, où S est un ensemble.

Alors, pour n'importe quel Q_0, Q_n va converger vers S presque surement.

Corollaire : F va décroire rapidement

Puisque $\Delta F = \mathbb{E}[F(Q(t+1)) - F(Q(t))|Q(t)] \leq 0$,

F va décroître presque surement.

Le théorème de Foster (Théorie des martingales) :
Soit $Q_0, Q_1, \cdots, Q_n, \cdots$ une suite de variables aléatoires ayant des valeurs dans E tel que pour certaine fonction $F : E \to \mathbb{R}^{\geq 0}$, $\mathbb{E}[F(Q_{n+1})|Q_n] \leq F(Q_n) - \epsilon$ chaque fois que $Q_n \notin S$, où S est un ensemble.

Alors, pour n'importe quel Q_0, Q_n va converger vers S presque surement.

De plus

$$\mathbb{E}[Reach(S)] \leq \frac{F(Q_0)}{\epsilon}$$

où $Reach(S)$ le temps où S est atteint pour la première fois.

Corollaire : Vitesse de convergence

- $\Delta F = -b\frac{1}{n}G$ où

$$G = \sum_{i=1}^{n} w_i \sum_{k=1}^{m} \sum_{k'>k} q_{i,k} q_{i,k'} [h_{i,k} - h_{i,k'}]^2 + \frac{1}{n}\mathcal{O}(b)$$

- Soit $\epsilon > 0$. Notons

$$Inf(\epsilon) = \{X | G(X) \leq \epsilon\}.$$

- Pour tout $\epsilon > 0$

$$\mathbb{E}[Reach(Inf(\epsilon))] \leq \frac{nF(Q(0))}{b\epsilon}$$

$$\mathbb{E}[Reach(Inf(\epsilon))] \leq \frac{3n(\sum_{j=1}^{N} w_j^2)(\sum_{k=1}^{m} 1/s_k)}{2b\epsilon}$$

Propriété 2 : Les points limites et les équilibres

- F décroît presque surement lorsque

$$\Delta F = -b\frac{1}{n}\sum_{i=1}^{n} w_i \sum_{k=1}^{m}\sum_{k'>k} q_{i,k}q_{i,k'}[h_{i,k} - h_{i,k'}]^2 + \frac{1}{n}\mathcal{O}(b) < 0$$

- Donc, ΔF doit s'approcher de 0 : un point où le premier terme s'annule correspond à un point tel que $\forall i \; h_{i,k} = h_{i,k'}$ $\forall k, k'$ tels que $q_{i,k} \neq 0$, $q_{i,k'} \neq 0$.

- Les points stables parmi ceux-ci correspondent à un équilibre de Nash.

4. Conclusion

Conclusion

- Nous avons considéré un algorithme d'apprentissage d'équilibres de Nash particulier.
- Il y a convergence presque sure de notre algorithme vers un équilibre de Nash pour les jeux d'ordonnancement.
- Il est possible de parler partiellement de la vitesse de convergence pour ces jeux.

La suite ?

Modification de l'algorithme

- Peut-on rendre l'algorithme polynomial : rendre le cardinal de l'ensemble des stratégies utilisées polynomial ?

Vitesse de convergence

- Vitesse de convergence dans le cas asynchrone ?
- Méthode de preuves pour le cas synchrone ?

D'autres dynamiques ?

- Condition(s) nécessaire(s) et suffisante(s) pour une convergence vers un équilibre de Nash.
- Liens avec la vitesse de convergence.

Merci de votre attention.

Les jeux de congestion

Retour

- Trois ressources $\{1,2,3\}$

  $C_1(x) = C_2(x) = x$ et $C_3(x) = 2.$
- Deux joueurs
  - Le joueur 1 a deux stratégies pures $\{1\}$ et $\{2\}$
  - Le joueur 2 a deux stratégies pures $\{1\}$ et $\{2,3\}$

		joueur 2	
		$\{1\}$	$\{2,3\}$
joueur 1	$\{1\}$	(2,2)	(1,3)
	$\{2\}$	(1,1)	(2,4)

Forme stratégique du jeu

		joueur 2	
		$\{1\}$	$\{2,3\}$
joueur 1	$\{1\}$	3	4
	$\{2\}$	2	4

Valeur de la fonction de potentiel

$\phi(Q) = \sum_{e\in[m]} \sum_{k=1}^{\lambda_e(Q)} c_e(k)$

avec $\lambda_e(Q)$ la charge de la ressource e