Troc Combinatoire à Monte-Carlo

Tristan Cazenave[1], Yann Chevaleyre[1], Gaétan Marceau, Nicolas Maudet[1]

LAMSADE, Univ. Paris-Dauphine
{FirstName.LastName}@dauphine.fr

Mots-Clés : *Partage de Ressources Indivisibles, Algorithmes de Monte-Carlo*

1 Allocation Distribuée de Ressources Indivisibles

Cet article s'intéresse au problème de partage de ressources indivisibles par des mécanismes distribués. Plus spécifiquement, on considère un ensemble $\mathcal{R}$ de ressources devant être attribuées à un ensemble $\mathcal{N}$ d'agents. Toutes les ressources doivent être distribuées et aucune n'est partageable ni divisible : une *allocation* est donc simplement une partition des ressources parmi les agents ($A : \mathcal{N} \to 2^{\mathcal{R}}$). Les agents ont des *préférences* sur les lots de ressources qu'ils peuvent posséder (typiquement selon les contextes une relation d'ordre sur les lots, ou une fonction d'utilité permettant de valuer les différents lots). Le problème général est d'allouer "optimalement" les ressources aux agents (typiquement d'une façon qui soit Pareto-optimale, ou qui maximise la somme des utilités des agents). A l'inverse des approches centralisées supposant l'existence d'un algorithme prenant en entrée les préférences des agents et calculant une allocation optimale, on cherche ici à concevoir le mécanisme de manière à ce qu'il exhibe des propriétés souhaitables en dépit du comportement auto-intéressé des agents. En particulier, on cherche à montrer la convergence de tels systèmes si les agents peuvent modifier successivement l'allocation courante en procédant à des échanges de ressources qu'ils trouvent *individuellement acceptables*, donnant ainsi lieu à des séquences d'échanges. A cet effet, on suppose que les agents utilisent un critère de rationalité simple : une transaction est inacceptable si elle baisse *directement* (on parle de critère myope) leur bien-être individuel. Le déroulement du processus de négociation peut à présent être décrit : partant d'une allocation initiale, les agents cherchent à s'accorder sur des transactions acceptables avec d'autres agents. Lorsqu'une transaction est acceptée par les agents impliqués elle est implémentée et permet de passer à une nouvelle allocation A', jusqu'à ce qu'aucune transaction ne puisse plus être réalisée. La négociation s'arrête alors. On voit que différents paramètres peuvent être instanciés : la complexité des transactions permises entre les agents (en particulier, on peut limiter le nombre de ressources ou le nombre d'agents impliqués dans les échanges), la nature des préférences des agents, la topologie du système, etc. Les propriétés de cette approche (convergence, complexité de communication, etc.) ont été étudiée en détail [**?**], mais essentiellement dans l'optique de caractériser des conditions permettant de garantir la convergence de tels processus.

Dans cet article, on s'attaque à une situation où aucune garantie de convergence ne peut être apportée. Pour permettre au système de dépasser cette limitation sans contredire le principe d'agents auto-intéressés, on propose ici d'étendre l'horizon de rationalité de ces agents en leur permettant d'échantilloner les déroulements possibles de la négociation selon le principe d'algorithmes de Monte-Carlo. Notre propos ici est de souligner les difficultés méthodologiques posées par cette approche et de donner de premiers résultats expérimentaux.

2 Horizon de Rationalité Etendu : le cas du Troc Combinatoire

Dans le contexte du troc combinatoire (pas d'utilisation de monnaie lors des transactions), il est intéressant de noter que si les transactions sont réduites à des *échanges bilatéraux simples* (deux agents échangent une ressource contre une autre ressource) même des restrictions très fortes sur la nature des préférences des agents ne permet pas de garantir de convergence du système vers une allocation Pareto-optimale. Considérons en effet l'exemple très simple suivant, où les utilités sont additives : l'agent a_1 value les objets de la manière suivante $r_1 \mapsto 1, r_2 \mapsto 2, r_3 \mapsto 3$, l'agent $a_2 : r_1 \mapsto 3, r_2 \mapsto 1, r_3 \mapsto 2$, et enfin l'agent $a_3 : r_1 \mapsto 2, r_2 \mapsto 3, r_3 \mapsto 1$. Supposons que l'allocation initiale soit $A(a_1) = \{r_2\}$, $A(a_2) = \{r_3\}$ et $A(a_3) = \{r_1\}$, ce qui donne un vecteur $\langle 2, 2, 2 \rangle$. Chaque agent possède sa deuxième ressource dans l'ordre de préférence. Cette allocation n'est *pas* Pareto-optimale (elle est évidemment strictement dominée par l'allocation qui donne à chacun sa ressource favorite), mais on constate qu'*aucune séquence d'échanges bilatéraux simples ne permet d'atteindre une situation Pareto-optimale*. Supposons pourtant que les agents puissent considérer les scénarii envisageables s'il acceptent une perte d'utilité à court terme. Prenons le cas de l'agent a_1 : en échangeant r_2 contre r_1 avec a_3 il accuse une perte d'utilité mais est *certain* de pouvoir échanger r_1 contre r_3 avec a_2 (c'est alors le seul échange possible). Souvent la situation ne sera pas si claire : certaines scenarii amèneront à des issues favorables, d'autres non. L'approche proposée ici est d'échantillonner les scenarii possibles selon le principe des algorithmes de Monte-Carlo afin de fonder la décision de certains agents. Les approches de Monte-Carlo ont récemment été utilisées avec succès dans de nombreux contextes, dont les problèmes d'échanges de reins entre patients [**?**]. Dans le cadre du troc combinatoire distribué proposé ici les difficultés sont nombreuses, en particulier : chaque agent est (potentiellement) opposé à de *nombreux* autres agents, qui peuvent interagir entre eux ; et l'allocation courante ainsi que les préférences des agents ne sont pas nécessairement complètement connues des agents. Nous distinguerons par la suite les agents de Monte-Carlo (qui disposent de la capacité de lancer un algorithme de Monte-Carlo avant de décider si une transaction est acceptable ou pas), des agents d'Arcachon, qui obéissent au critère myope habituel. Le paramétrage des scenarii aléatoires à considérer n'est pas sans poser problème. On peut distinguer :

1. *la longueur et le nombre de scenarii*. Ce sont les paramètres classiques de ce type d'algorithmes (la longueur correspond ici aux nombres d'échanges qu'un agent simulera en avant).
2. *le périmètre de la négociation* : on distingue les scenarii aléatoires *focalisés* où les agents n'envisagent que les transactions qui les concernent, des scenarii *non focalisés* où les agents considèrent également les échanges qui peuvent intervenir entre d'autres agents ;
3. *le niveau d'informations des agents* : ce paramètre peut affecter différents aspects. Quelle est la connaissance que les agents ont de *l'allocation courante* ? Quelle est la connaissance que les agents ont des *préférences* des autres agents ? Quelle est la connaissance que les agents ont des *critères de rationalité* et des *stratégies* utilisées par les autres ?

On a mené de premières expérimentations pour évaluer la pertinence et la dépendance de certains de ces paramètres dans notre contexte. Le type de résultat finalement attendus sont : (i) une société d'agents de Monte-Carlo se comporte-t-elle de manière plus satisfaisante qu'une société d'Arcachon ? (ii) un agent de Monte-Carlo se comporte-t-il mieux que la moyenne au sein d'une société d'Arcachon ?

Références

[1] P. Awasthi and T. Sandholm. Online stochastic optimization in the large : Application to kidney exchange. In *Proceedings of IJCAI*, 2009.

[2] Y. Chevaleyre, U. Endriss, and N. Maudet. Simple negotiation schemes for agents with simple preferences : Sufficiency, necessity and maximality. *Journal of Autonomous Agents and Multiagent Systems*, 2010. In press.