Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence

Symmetric Subgame Perfect Equilibria in Resource Allocation

Ludek Cigler
Ecole Polytechnique Fédérale de Lausanne
Artificial Intelligence Laboratory
CH-1015 Lausanne, Switzerland
ludek.cigler@epfl.ch

Boi Faltings
Ecole Polytechnique Fédérale de Lausanne
Artificial Intelligence Laboratory
CH-1015 Lausanne, Switzerland
boi.faltings@epfl.ch

Abstract

We analyze symmetric protocols to rationally coordinate on an asymmetric, efficient allocation in an infinitely repeated N-agent, C-resource allocation problems. (Bhaskar 2000) proposed one way to achieve this in 2-agent, 1-resource allocation games: Agents start by symmetrically randomizing their actions, and as soon as they each choose different actions, they start to follow a potentially asymmetric "convention" that prescribes their actions from then on. We extend the concept of convention to the general case of infinitely repeated resource allocation games with N agents and C resources. We show that for any convention, there exists a symmetric subgame perfect equilibrium which implements it. We present two conventions: bourgeois, where agents stick to the first allocation; and market, where agents pay for the use of resources, and observe a global coordination signal which allows them to alternate between different allocations. We define price of anonymity of a convention as the ratio between the maximum social payoff of any (asymmetric) strategy profile and the expected social payoff of the convention. We show that while the price of anonymity of the bourgeois convention is infinite, the market convention decreases this price by reducing the conflict between the agents.

1 Introduction

In many situations, agents have to coordinate their use of some resource. One wireless channel can only be used by one device, one parking slot may only be occupied by one vehicle, etc. The problem is that often, the agents have identical preferences: Everyone prefers to access rather than yield. Similarly, everyone prefers a parking slot closest to the building entrance. However, if multiple agents try to use one resource simultaneously, they collide and everyone loses.

Consider a simple example: two agents who want to access a single resource. We can describe the problem as a game. Both agents have two actions: yield (Y) and access (A). If agent α yields, it gets a payoff of 0. When agent α accesses the resource while the other agent yields, it gets a payoff of 1. But if both agents access the resource at the same time, they both incur a cost $\gamma < 0$.

The normal form of such a game looks as follows:

	Y	A
Y	$0, 0$	$0, 1$
A	$1, 0$	γ, γ

This is a symmetric game, but the two efficient Nash equilibria (NE) are asymmetric: either one agent yields and the other one accesses the resource, or vice versa. The only symmetric rational outcome is the mixed NE where both agents access the resource with probability $\Pr(A) := \frac{1}{|\gamma|+1}$. However, this mixed equilibrium is not efficient, because the expected payoff of both agents get is 0.

Asymmetric equilibria of symmetric games are undesirable for two reasons: First, they are not fair. In our example, only one agent can access the resource. Second, coordinating on an asymmetric equilibrium is difficult. Imagine that the agents are all identical and anonymous, i.e. they cannot observe neither their own identity, neither the identity of any other agent. We cannot prescribe a different strategy for each of the agents.

In our previous work (Cigler and Faltings 2011), we addressed the fairness issue. We considered a special case of a resource allocation problem. We proposed to use a global coordination signal and multiagent learning to reach a symmetric, fair and efficient wireless channel allocation ((Wang et al. 2011) later implemented this approach in an actual wireless network and achieved throughput 3x higher than standard ALOHA protocols). The advantage is that the coordination signal is not specific to the game. It does not have to tell the agents which action to take. As an example, the agents may just observe noise on some frequency. The disadvantage of our previous approach though was that it was not rational for the agents to adopt this algorithm – an agent who decided to always access the resource could force everyone else out and achieve the highest payoff.

In this paper, we consider the rationality issue. We propose a distributed algorithm to find an allocation of a set of resources which is not only symmetric and fair, but also rational. We draw inspiration from the works of (Bhaskar 2000) and (Kuzmics, Palfrey, and Rogers 2010) on symmetric equilibria for symmetric repeated games.

Assume that agents play an infinitely repeated game, and they discount future payoffs with a common discount factor $0 < \delta < 1$. A strategy for an agent is a mapping from any history of the play to a probability distribution over the actions. Our goal is to find a symmetric *subgame perfect equilibrium*. A subgame perfect equilibrium is a strategy profile (vector of strategies for every agent) which is a NE in any history, including those that cannot occur on the equilibrium path.

The problem is that the agents start with a common history. In order to ever use different actions, they need to randomize; in order to randomize, they need to be *indifferent* between a set of actions.

We can split the game in two stages: Symmetric play, when all the agents use the same actions, asymmetric play from then on. We call the first round of the game where agents differ an *asynchrony*.

After reaching asynchrony, agents can proceed in different ways, depending on which action they took in the asynchrony round. We call the strategy profile that the players adopt after an asynchrony a *convention*. The agents who have successfully accessed a resource alone in that round are "winners", and all the other agents are "losers". The convention can prescribe a different strategy for the winners and for the losers.

As an example, for the 2-agent, 1-resource allocation game, (Bhaskar 2000) describes the following two conventions:

Bourgeois Agents keep using the action they played in the last round;

Egalitarian Agents play the action of their opponent from the last round.

Some form of convention is necessary to achieve and maintain coordination even in games where agents don't have conflicting preferences ((Crawford and Haller 1990), (Goyal and Janssen 1996)). We will assume that a convention is rational, i.e. every agent is playing their best response actions.

The social payoff depends on how fast the agents reach asynchrony. When there is a big difference between the winner and loser payoff, the losers will "fight back" harder, so they will play their most preferred action with higher probability. In the egalitarian convention, the payoffs to the loser are (given high enough δ) the same as the winner. Therefore, the agents will be indifferent between being a winner and a loser and they will reach asynchrony faster.

How much social payoff do we lose by requiring the agents to be identical and anonymous, and requiring them to play only symmetrical strategies? For a given convention, we can calculate the social payoff E the agents get if we implement it as a symmetric SPE. That way, we can define its *price of anonymity*. It is the ratio between E and the highest expected social payoff of any (potentially asymmetric) strategy profile.

Definition 1. *Let $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_N)$ be a symmetric strategy vector for the resource allocation game of N agents and C resources. We define the* price of anonymity of strategy vector σ *as follows:*

$$R(\sigma) := \frac{\max E(\tau)}{E(\sigma)}$$

where $E(\sigma)$ is the social expected payoff (sum of individual expected payoffs) when agents adopt strategy σ, and $\max E(\tau)$ is the maximum social payoff of any strategy vector, symmetric or asymmetric.

For a given resource allocation game, we can also define its price of anonymity *as*

$$R := \inf R(\sigma)$$

where we minimize over all symmetric strategy profiles for the given game.

Our main contributions are the following:

- We show that in the infinitely repeated resource allocation game with discounting of N agents and C resources, for any convention, there exists a symmetric subgame perfect equilibrium that reaches this convention.
- We give a closed form expression to calculate the subgame perfect equilibrium for the bourgeois convention, and show that for a small number of resources C, this convention leads to zero expected payoff. This means that the price of anonymity of the bourgeois convention is ∞.
- We present the market convention. It is based on the idea that agents can observe a common coordination signal, and they can reach a different resource allocation for each signal. We show that compared to the bourgeois convention, it improves the expected payoff. Its price of anonymity is therefore finite.

This paper is structured as follows: In Section 2 we present our resource allocation game. We show that for any convention there exists a symmetric subgame perfect equilibrium which implements this convention. In Section 3 we present two concrete examples of a convention: *bourgeois* and *market* conventions and discuss their properties. Finally, Section 4 concludes.

2 Resource Allocation Game

Definition 2. *A resource allocation game $\mathcal{G}_{N,C}$ is a game of N agents. Each agent i can access one of C resources. It chooses its action a_i from $\mathcal{A}_i = \{Y, A_1, A_2, \ldots, A_C\}$, where action $a_i = Y$ means to yield, and action $a_i = A_c$ means to access resource c. Because all resources are identical, we can define a special meta-action $a_i = A$. To take action A means to choose to access, and then to choose the resource uniformly at random.*

The payoff function for agent i is defined as follows:

$$u_i(a_1, \ldots, a_i, \ldots, a_N) := 0 \text{ if } a_i = Y \quad (1)$$

$$u_i(a_1, \ldots, a_i, \ldots, a_N) := \begin{cases} 1 & \text{if } a_i \neq Y, \\ & \forall j \neq i, a_j \neq a_i \\ \gamma < 0 & \text{otherwise} \end{cases} \quad (2)$$

This game has a set of pure strategy NEs where C agents each access a resource c_i and $N - C$ agents do not. There is also a symmetric mixed strategy NE in which each agent decides to access *some* resource with probability

$$\Pr(a_i > 0) := \min\left\{C \cdot \left(1 - \sqrt[N-1]{\frac{|\gamma|}{1+|\gamma|}}\right), 1\right\} \quad (3)$$

and then chooses the resource to access uniformly at random. Note that for high enough values of C, all agents will choose to access *some* resource.

As mentioned before, the pure strategy NEs are efficient, but neither symmetric nor fair. For a small number of resources C, the mixed strategy NE leads to an expected payoff of 0. Therefore, it is not socially efficient.

To improve the efficiency, we will follow the approach of Bhaskar and Kuzmics et al. (see Section 1). We consider an infinitely repeated version of the resource allocation game $\mathcal{G}_{N,C}$. We assume that the agents discount their future payoffs with a discount factor $0 < \delta < 1$. Agents have information about which resources have been occupied in the last round, but they cannot observe *who* occupied them.

When N identical agents use symmetric strategies to play the repeated game, in order to reach an asymmetric outcome of a single shot game they need to distinguish themselves. One way to do this is as follows: Whenever an agent is the only one to access a resource, he becomes the "winner". The other agents (those who yielded or collided) are "losers". We call this event *asynchrony*.

The problem is that there can only be as many winners as there are resources. We won't distinguish between the losers. Another way to distinguish them is as follows: We assume that the agents can observe in each round of the game a global coordination signal, which is an integer $k \in \{1, 2, \ldots, K\}$, chosen uniformly at random. This signal is the same for all the agents. They can condition their strategy on this signal. For different coordination signals, there can be different sets of winners and losers.

Definition 3. *A* strategy *σ of an agent α is a function from the coordination signal to a probability distribution over actions,*

$$\sigma : \{1, 2, \ldots, K\} \to \Delta(\{A, Y\}) \tag{4}$$

A deterministic strategy *selects for each signal either A, or Y.*

Definition 4. *Let r be a round of the infinitely repeated game $\mathcal{G}_{N,C}$. We say round r is* an asynchrony*, if there exists an agent α who:*

1. *Accesses some resource alone in round r,*
2. *and has not accessed any resource alone in previous rounds.*

We call all such agents winners*. All the other agents who have not accessed any resource alone so far are* losers.

Definition 5. *A* signal-based convention *(or simply* convention*) ξ in a game $\mathcal{G}_{N,C}$ is a set of (mixed) strategies that the agents adopt after an* asynchrony *round. The strategies for the winners and for the losers are potentially different. Suppose there are n_w winners. The convention leads to an expected payoff $w_\xi(n_w)$ for the winners, and $l_\xi(n_w)$ for the losers.*

Figure 1 gives an example of a game play of $N = 4$ agents, $C = 2$ resources and $K = 2$ signals. If in round t, the agents observe a signal k_t, the convention adopted by the agents in this example prescribes that if an agent accesses a resource alone in round t, it becomes its "winner" and will

Round	**1**	2	**3**	**4**	5	6	**7**	8
Signal	1	2	2	2	1	2	1	2
Agent 1	1	0	2	0	1	0	**1**	0
Agent 2	1	0	**1**	1	1	1	0	1
Agent 3	**2**	1	0	0	2	0	2	0
Agent 4	0	1	2	**2**	1	2	0	2

Figure 1: Example of a game play for $N = 4$ agents, $C = 2$ and $K = 2$. The asynchrony rounds are 1, 3, 4 and 7 (denoted in bold face). Once an agent accesses a resource alone, it will keep accessing that resource every time the same signal is observed. The winners are denoted with gray background (different shades for different signals). The first round when an agent accesses a resource alone (and becomes the winner) is denoted with bold face. In the rest of the game, agent 1 will keep accessing resource 1 when the signal is 1. Agent 2 will access the resource 1 when signal is 2. Agent 3 will access the resource 2 when the signal is 1. Finally, agent 4 will access the resource 2 when the signal is 2. This way, the agents are no longer anonymous and have identified their roles with the signal and the resource they access.

access the same resource in every round $t' > t$ in which the signal $k_{t'} = k_t$.

Definition 6. *Let ξ be a convention for game $\mathcal{G}_{N,C}$. Let σ be a deterministic strategy of an agent α. Assume that for each signal $k \in \{1, \ldots, K\}$, every other agent takes action A with probability p_k. Let $\vec{p} = (p_1, p_2, \ldots, p_K)$ be a vector of these probabilities. We define* expected payoff functions E_A and E_Y *when agent α takes actions A and Y:*

$$\begin{aligned} E_A(\vec{p}, \sigma, k) &:= \sum_{c=1}^{C} [\Pr(\alpha \textit{ wins } \& \; n_w = c | A) w_\xi(c) \\ &+ \Pr(\alpha \textit{ loses } \& \; n_w = c | A)(\gamma + l_\xi(c))] + \Pr(n_w = 0 | A) \\ &\cdot \left[\gamma + \frac{\delta}{K}\left(E_A(\vec{p}, \sigma, k) + \sum_{\substack{l=1 \\ l \neq k}}^{K} E_{\sigma(l)}(\vec{p}, \sigma, l)\right)\right] \end{aligned} \tag{5}$$

$$\begin{aligned} E_Y(\vec{p}, \sigma, k) &:= \sum_{c=1}^{C} \Pr(n_w = c | Y) \cdot l_\xi(c) \\ &+ \Pr(n_w = 0 | Y) \frac{\delta}{K}\left(E_Y(\vec{p}, \sigma, k) + \sum_{\substack{l=1 \\ l \neq k}}^{K} E_{\sigma(l)}(\vec{p}, \sigma, l)\right) \end{aligned} \tag{6}$$

Lemma 1. *For any strategy σ and signal k, the functions E_A and E_Y are continuous in $\vec{p} \in \langle 0, 1\rangle^K$.*

Proof. The probabilities $\Pr(n_w = c|A)$ and $\Pr(n_w = c|Y)$ are continuous. The functions E_A and E_Y are sums of products of continuous functions, so they must be themselves continuous. □

Lemma 2. *Functions E_A and E_Y are well-defined for any σ, k and $\vec{p} \in \langle 0,1 \rangle^K$.*

Proof. For fixed $\vec{p}$, σ, γ and δ the functions E_A, E_Y define each a system of K linear equations. We can write this system as $\mathbf{A}\vec{E}_\sigma = \mathbf{b}$, where $\vec{E}_\sigma$ is a vector of corresponding payoff functions $E_{\sigma(k)}$, and $\mathbf{b} \in \mathbb{R}^K$. The matrix $\mathbf{A}$ is defined as

$$\mathbf{A} := \mathbf{I} - \frac{\delta}{K}(\Pr(n_w = 0|\sigma(1)), \dots, \Pr(n_w = 0|\sigma(K))) \cdot \mathbf{1}^T \tag{7}$$

where $\mathbf{I}$ is a $K \times K$ unit matrix and $\mathbf{1}^T$ a K-dimensional row vector of all 1.

This system of equations has a unique solution if the matrix $\mathbf{A}$ is non-singular. This is equivalent to saying that $\det(\mathbf{A}) \neq 0$.

The matrix $\mathbf{A}$ is diagonally dominant, that is $a_{ii} > \sum_{j=1, j \neq i}^{K} |a_{ij}|$. This is because $0 < \delta < 1$, and all the probabilities $\Pr(n_w = c|\sigma(k)) \leq 1$. It is known that diagonally dominant matrices are non-singular ((Taussky 1949)). Therefore, a unique solution $\vec{E}_\sigma$ of the system exists and the functions E_A, E_Y are well-defined. □

Suppose that given the probability vector $\vec{p}$, there is a deterministic best-response strategy for agent α $\sigma_{\vec{p}}$.

Theorem 1. *If the functions $E_A(\vec{p}, \sigma_{\vec{p}}, k)$ and $E_Y(\vec{p}, \sigma_{\vec{p}}, k)$ are well-defined and continuous in any p_k, there exists a probability vector $\vec{p}^* = (p_1^*, p_2^*, \dots, p_K^*)$ such that when for signal k, every agent accesses a resource with probability p_k^*, agents play a symmetric subgame perfect equilibrium of the infinitely repeated resource allocation game.*

Proof. Fix γ, δ, σ and $\vec{p}$ for all $l \in \{1, \dots, K\}$, $l \neq k$.

Let $p_k = 0$. If $E_Y \geq E_A$, everyone is best off playing Y and it is a symmetric best-response.

If not, then let $p_k = 1$. If in this case $E_A \geq E_Y$, everyone is best off playing A and again this is a symmetric best-response.

Finally, if both $E_Y < E_A$ for $p_k = 0$, and $E_Y > E_A$ for $p_k = 1$, then from the fact that both functions are well-defined and continuous for $0 \leq p_k \leq 1$, they must intersect for some $0 < p_k^* < 1$. For such p_k^*, the agents are indifferent between actions A and Y. Therefore, it is a symmetric best-response when all agents play A with probability p_k^*.

We now know that for any coordination signal k, there exists a symmetric best-response given any set strategies $\sigma(l)$ for other coordination signals $l \neq k$. Therefore, there must exist a probability vector $\vec{p}^*$ such that for all coordination signals it is a symmetric best-response to access with these probabilities. This $\vec{p}^*$ defines a symmetric subgame perfect equilibrium of the infinitely repeated game. □

2.1 Calculating the Equilibrium

While the symmetric subgame perfect equilibrium is guaranteed to exist, in order to actually play it, the agents need to be able to calculate it. It is not always possible to obtain the closed form of the probability of accessing a resource. Therefore, we will show how to calculate the equilibrium strategy numerically.

Let $\vec{p}$ be a probability vector, σ a strategy and k a signal. Let $\vec{p}_0 := (p_1, p_2, \dots, p_k = 0, \dots, p_K)$, i.e. vector $\vec{p}$ with p_k set to 0. Let $\vec{p}_1 := (p_1, p_2, \dots, p_k = 1, \dots, p_K)$. From Theorem 1 we know that either $E_Y(\vec{p}_0, \sigma, k) > E_A(\vec{p}_0, \sigma, k)$, or $E_A(\vec{p}_1, \sigma, k) > E_Y(\vec{p}_1, \sigma, k)$ or the two functions intersect for some $0 \leq p_k \leq 1$. Furthermore, we know that $E_A(\vec{p}_0, \sigma, k) = w_\xi(c)$ since the probability of successfully claiming a resource is 1 when everyone else yields, and also $E_Y(\vec{p}_0, \sigma, k) = 0$. Therefore, $E_Y(\vec{p}_0, \sigma, k) > E_A(\vec{p}_0, \sigma, k)$ *iff* $w_\xi(c) > 0$.

W.l.o.g, we will assume that $w_\xi(c) > 0$. Algorithm 1 shows then how to calculate the probability vector.

Algorithm 1 Calculating the equilibrium probabilities

for Each subset $S \subseteq \{1, 2, \dots, K\}$ **do**
  Let Σ be a system of equations
  $\forall i \notin S$, Σ contains two equations for $E(\vec{p}, \sigma, i)$. One corresponding to $E_A(\vec{p}, \sigma, i)$, one to $E_Y(\vec{p}, \sigma, i)$.
  $\forall j \in S$, we set $p_j := 1$. Σ contains only one equation for $E(\vec{p}, \sigma, j)$, corresponding to $E_A(\vec{p}, \sigma, j)$.
  So Σ is a system of $2K - |S|$ equations with $2K - |S|$ variables.
  Solve numerically the system of equations Σ.

  if there exists a solution to Σ for which $\forall i \notin S$, $0 \leq p_i \leq 1$ **then**
    We have found a solution
    break;
  end if
end for

3 Conventions

In the previous section, we have shown that we can find a symmetric way to reach any convention, provided the agents access the resources with a certain probability. We have also shown how to calculate the resource access probability in every stage of the game. In this section, we would like to show specific examples of the conventions that agents can adopt, and discuss their properties.

3.1 Bourgeois Convention

The bourgeois convention is the simplest one. Once an agent has accessed a resource successfully for the first time, he will keep accessing it forever. We say that the agent has *claimed* the resource. We don't need any coordination signal to implement it, so we assume that $K = 1$.

For N agents and C resources, we will describe the decision problem from the point of view of agent α. Let c be the number of resources that have not been claimed yet, and $n := N - C + c$ the number of agents who have not claimed a resource yet. We define $E(c, \tau_{-\alpha})$ as the expected payoff of the best response strategy for agent α given the strategies $\tau_{-\alpha}$ of all the opponents.

Lemma 3. *For any $\tau_{-\alpha}$ and $\forall c \geq 1$, $E(c, \tau_{-\alpha}) \geq 0$.*

Proof. No matter what is the strategy of the opponents, if agent α chooses to always yield, its payoff will be 0. □

Lemma 4. *If the opponents' strategies $\tau_{-\alpha}$ are such that the agent α is indifferent in every round between yielding and accessing, $E(c, \tau_{-\alpha}) = 0$ for all $c \geq 1$.*

Proof. If the agent α is indifferent between actions Y and A in every round, that means that it is indifferent between a strategy that prescribes Y in every round and any other strategy. The (expected) payoff of the strategy which prescribes always Y is 0. Therefore, the expected payoff of any other strategy must be 0 as well. □

For the purpose of our problem, all the unclaimed resources are identical. Therefore the only parameter of the agent strategy is the probability with which it decides to access – the resource itself is then chosen uniformly at random. Lemma 4 shows a necessary condition for agent α to be indifferent. The following lemma shows a sufficient condition:

Lemma 5. *Assume at round r there are c unclaimed resources. Then there exists a unique $0 \leq p^* \leq c$ such that if all opponents who haven't claimed any resource yet play A with probability $p^*_c = c\left(1 - \sqrt[n-1]{\frac{|\gamma|}{|\gamma| + \frac{1}{1-\delta}}}\right)$, agent α is indifferent between yielding and accessing.*

Proof. From Lemma 4 we know that when agent α is indifferent, it must be that $E(c, \tau_{-\alpha}) = 0$ for all $c \geq 1$.

The expected profit to agent α from playing A and then following best-response strategy (with zero payoff) is

$$E_A(c, \tau_{-\alpha}) = \left(1 - \frac{p}{c}\right)^{n-1} \cdot \frac{1}{1-\delta} + \left[1 - \left(1 - \frac{p}{c}\right)^{n-1}\right] \cdot \gamma \tag{8}$$

Here p is the probability with which the opponents access. We want $E_A(c, \tau_{-\alpha}) = E_Y(c, \tau_{-\alpha}) = 0$. This holds if p^*_c is defined as in the theorem above.

Function E_A is decreasing in p on the interval $[0, c]$, while function E_Y is constantly 0. Therefore, the intersection is unique on an interval $[0, c]$. □

Lemma 6. *Assume that all the opponents who haven't claimed any resource access a resource with probability $p < p^*_c$. Then it is best-response for agent α to access.*

Proof. The probability that agent α claims successfully a resource after playing A is

$$\Pr(\text{claim some resource}|A) := \left(1 - \frac{p}{c}\right)^{n-1} \tag{9}$$

This probability increases as p decreases. Therefore the expected profit of accessing is increasing, whereas the profit of yielding stays 0. □

Theorem 2. *Define an agent's strategy τ as follows: If there are c unclaimed resources, play A with probability $p_c := min\left(1, p^*_c\right)$ (where p^*_c is defined in Lemma 5). Then a joint strategy profile $\vec{\tau} = (\tau_1, \tau_2, \ldots, \tau_N)$ where $\forall c, \tau_c = \tau$ is a subgame perfect equilibrium of the infinitely repeated resource allocation game.*

Proof. If $p^*_c < 1$, any agent is indifferent between playing Y and playing A, therefore will happily follow strategy τ. If $1 = p_c < p^*_c$, it is best response for any agent to play A, just as the strategy τ prescribes. □

Theorem 3. *For all $c \in \mathbb{N}$, if $p_c = p^*_c$, $E(c, \vec{\tau}_{-\alpha}) = 0$.*

Proof. We will proceed by induction.

For $c = 0$, the expected payoff is trivially $E(0, \vec{\tau}_{-\alpha}) = 0$, because there are no free resources.

Let $\forall j < c, E(j, \vec{\tau}_{-\alpha}) = 0$ and $p_c = p^*_c$. If agent α plays Y, the expected payoff is clearly 0 (it will be 0 now and 0 in the future from the induction hypothesis). If agent α plays A, the expected payoff is

$$\begin{aligned} E_A(c, \vec{\tau}_{-\alpha}) := &\left(1 - \frac{p_c}{c}\right)^{n-1} \cdot \frac{1}{1-\delta} \\ &+ \left[1 - \left(1 - \frac{p_c}{c}\right)^{n-1}\right] \cdot \gamma + \delta \sum_{j=0}^{c} q_{cj} E(j) \end{aligned} \tag{10}$$

Because of the way the p^*_c is defined, and from the induction hypothesis $E(j, \vec{\tau}_{-\alpha}) = 0$ for $j < c$, we get

$$\begin{aligned} E_A(c, \vec{\tau}_{-\alpha}) &:= \delta q_{cc} E(c, \vec{\tau}_{-\alpha}) \\ &= \delta q_{cc} \max\{E_A(c, \vec{\tau}_{-\alpha}), E_Y(c, \vec{\tau}_{-\alpha})\} \end{aligned} \tag{11}$$

Since $\delta q_{cc} < 1$, it must be that $E_A(c, \vec{\tau}_{-\alpha}) = 0$. □

Theorem 4. *If $p_c < p^*_c$, $E(c, \vec{\tau}_{-\alpha}) > 0$.*

Proof. From Lemma 6 we know that when $p_c < p^*_c$, it is a best response to access, so $E(c, \vec{\tau}_{-\alpha}) = E_A(c, \vec{\tau}_{-\alpha})$. From Lemma 3 we know that for all j, $E(j) \geq 0$. If $p_c < p^*_c$, from the definition of $E_A(c, \vec{\tau}_{-\alpha})$ (Equation 10) we see that $E(c, \vec{\tau}_{-\alpha}) > 0$. □

Theorem 4 shows that if we have enough resources so that $p^*_c \geq 1$, the expected payoff for the agents, even when they access all the time, will be positive.

Let us now look at the price of anonymity for the bourgeois convention (as defined in Definition 1). The highest social payoff any strategy profile τ can achieve in an N-agent, C-resource allocation game ($N \geq C$) is

$$\max E(\tau) := \frac{C}{1-\delta}. \tag{12}$$

This is achieved when in every round, every resource is accessed by exactly one agent. Such strategy profile is obviously asymmetric.

If each agent knew which part of the bourgeois convention to play at the beginning of the game, this convention would be socially efficient. However, when the agents are anonymous, they have to learn which part of the convention they should play through randomization. For the bourgeois convention (when C is small), this randomization wipes out all the efficiency gains. Therefore, its price of anonymity is infinite.

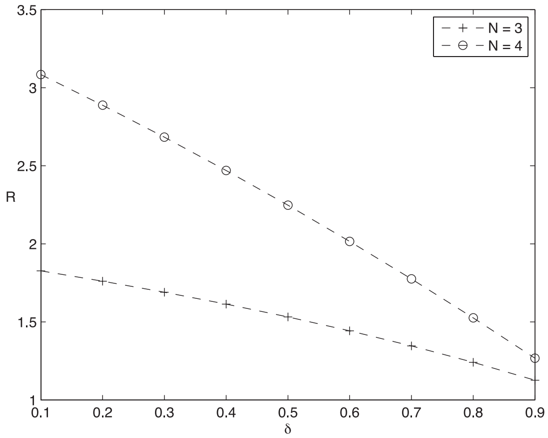

Figure 2: Market convention: Price of anonymity for $\gamma = -0.5$ and varying δ.

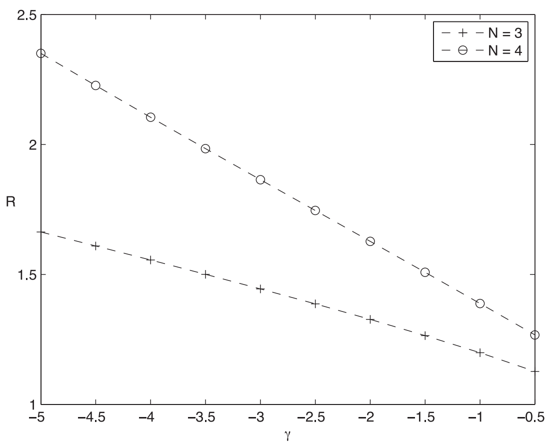

Figure 3: Market convention: Price of anonymity for $\delta = 0.9$ and varying γ.

3.2 Market Convention

We saw that the bourgeois convention leads to zero expected payoff for a small number of resources. We would like to improve the expected payoff here.

We assume the following:

- Agents can observe $K \geq 1$ coordination signals.
- Agents have a decreasing marginal utility when they access a resource more often.
- They pay a fixed price per each successful access, to the point that each agent prefers to access a resource only for one signal out of K. In practice, this could be implemented by a central authority which observes the convergence rate of the agents, and dynamically increases or decreases the price to achieve convergence.

Such assumptions define what we call "market" convention, in which the winners only access their claimed resource for the signals they observed when they first claimed it.

We know that we can implement this convention for $C \geq 1$ resources using symmetric play (see Section 2). We can

	ex-post fair	efficient	rational
C&F'11	(✓)[1]	✓	*no*
Bourgeois	*no*	*no*	✓
Egalitarian[2]	✓	✓	✓
Market	✓	?	✓

Table 1: Properties of conventions

also use Algorithm 1 to calculate the access probabilities. Here we will look specifically at the expected payoff of the market convention in the case of N agents and 1 resource.

When each agent only accesses the resource for one signal, we need $K = N$ signals to make sure everyone gets to access once.

In the N-agent, 1-resource case, imagine there are still n agents playing and $(N-n)$ agents who have already claimed the resource for some signal. Imagine that the n agents observe one of the n signals for which no resource has been claimed.

Assume that all agents access the resource with probability p_n. The expected payoff of accessing a resource for agent α is

$$E_A(p_n, n) := (1-p_n)^{n-1} \cdot \left(1 + \frac{\delta}{N} \cdot \frac{1}{1-\delta}\right) + \left[1-(1-p_n)^{n-1}\right] \cdot \left[\gamma + \frac{\delta n}{N-\delta(N-n)} E_A(p_n, n)\right] \tag{13}$$

The expected payoff of yielding for agent α is

$$E_Y(p_n, n) := (n-1)p_n(1-p_n)^{n-2} E(n-1) + \left[1-(n-1)p_n(1-p_n)^{n-2}\right] \frac{\delta n}{N-\delta(N-n)} E_Y(p_n, n) \tag{14}$$

When $p_n = 1$, accessing a resource will always lead to a collision, so the payoff will be negative. When $p_n = 0$, accessing a resource will always claim it, so the payoff will be positive. So in the equilibrium, the agents should be indifferent between accessing and yielding. Therefore, we want to find p_n^* such that $E_A(p_n^*, n) = E_Y(p_n^*, n) = E(n)$.

Finding a closed form expression for p_n^* is difficult, but we can use Algorithm 1 to calculate this probability, as well as the expected payoff $E(n)$, numerically.

Figures 2 and 3 show the price of anonymity of the market convention (as defined in Definition 1) of the market convention for varying discount factor δ, and varying cost of collision γ, respectively. From Section 3.1, we saw that the price of anonymity for $C = 1$ is ∞. On the contrary, for the market convention this price is in both cases finite and relatively small.

3.3 Convention Properties

We compare the properties of the following conventions: *C&F'11*, a channel allocation algorithm presented in (Cigler

[1] Fair asymptotically, as $N \to \infty$.

[2] Only for 2-agents, 1-resource games.

and Faltings 2011); *bourgeois*, presented in Section 3; *egalitarian*, presented in Section 1; and *market*, presented in this work.

We compare the conventions according to the following properties:

Ex-post fairness Is the expected payoff to all agents the same *even after asynchrony*?

Efficiency Does the convention maximize social welfare among all possible conventions?

Rationality Is it an equilibrium for the agents to adopt the convention?

Table 1 summarizes the properties of the conventions. The *C&F'11* convention is only approximately ex-post fair. The fairness is improving as the number of coordination signals increases, but some agents might have a worse payoff than others. On the other hand, it is efficient, at least with no discounting ($\delta = 1$). However, it is not rational. The bourgeois convention is neither fair nor efficient, in fact the expected payoff to the agents is 0 (for a small number of resources). It is rational though, since the agents are indifferent between being a winner and a loser. The egalitarian convention is fair, efficient and rational. However, it only works for games of 2 agents and 1 resource. Finally, the market convention is fair and rational. It is clearly more efficient than the bourgeois convention. Nevertheless, finding the most efficient convention remains an open problem.

4 Conclusions

In this paper, we considered the problem of equilibrium selection in the infinitely repeated resource allocation game with discounting of N agents and C resources. We assumed that the agents are identical, and that they use symmetric strategies. We based our work on the idea of (Bhaskar 2000): we let the agents play a symmetric mixed strategy, after which they adopt a certain convention. We show that for any convention, there exists a symmetric subgame perfect equilibrium that implements it. We presented two such conventions for the repeated resource allocation game: bourgeois and market convention. We defined the price of anonymity as the ratio between the expected social payoff of the best asymmetric strategy profile and the expected social payoff of a given symmetric strategy profile. We showed that while the price of anonymity for the bourgeois convention is infinite (at least for small number of resources), the price of anonymity of the market convention is finite and relatively small.

In the future work, we would like to investigate whether there exist more efficient conventions than the market convention (i.e. conventions with smaller price of anonymity). In general, finding an optimal convention is an NP-hard problem (Balan, Richards, and Luke 2011), but for a more restricted set of infinitely repeated resource allocation games, we might be able to find the optimal convention, similar to the Thue-Morse sequence (Richman 2001) used by (Kuzmics, Palfrey, and Rogers 2010) in the Nash demand game.

Acknowledgements

We are particularly thankful to David Parkes for giving the first author the unique opportunity to spend a few weeks in his lab at Harvard. David's openness and unparalleled knowledge is what really helped to originate this work. We would also like to thank Kate Larson for reading the draft of this paper and helping make the theoretical analysis much more readable.

References

Balan, G.; Richards, D.; and Luke, S. 2011. Long-term fairness with bounded worst-case losses. *Autonomous Agents and Multi-Agent Systems* 22(1):43–63.

Bhaskar, V. 2000. Egalitarianism and efficiency in repeated symmetric games. *Games and Economic Behavior* 32(2):247–262.

Cigler, L., and Faltings, B. 2011. Reaching correlated equilibria through multi-agent learning. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, 509–516. International Foundation for Autonomous Agents and Multiagent Systems.

Crawford, V. P., and Haller, H. 1990. Learning how to cooperate: Optimal play in repeated coordination games. *Econometrica* 58(3):571–595.

Goyal, S., and Janssen, M. 1996. Can we rationally learn to coordinate? *Theory and Decision* 40(1):29–49.

Kuzmics, C.; Palfrey, T.; and Rogers, B. 2010. Symmetric players in repeated games: Theory and evidence.

Richman, R. 2001. Recursive binary sequences of differences. *Complex Systems* 13(4):381–392.

Taussky, O. 1949. A recurring theorem on determinants. *The American Mathematical Monthly* 56(10):672–676.

Wang, L.; Wu, K.; Hamdi, M.; and Ni, L. M. 2011. Attachment learning for multi-channel allocation in distributed OFDMA networks. *Parallel and Distributed Systems, International Conference on* 0:520–527.