Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)

Equilibria Under the Probabilistic Serial Rule

Haris Aziz and **Serge Gaspers** and **Simon Mackenzie** and **Nicholas Mattei**
NICTA and UNSW, Sydney, Australia
{haris.aziz, serge.gaspers, simon.mackenzie, nicholas.mattei}@nicta.com.au

Nina Narodytska
Carnegie Mellon University, Pittsburgh, USA
ninan@cs.cmu.edu

Toby Walsh
NICTA and UNSW, Sydney, Australia
toby.walsh@nicta.com.au

Abstract

The probabilistic serial (PS) rule is a prominent randomized rule for assigning indivisible goods to agents. Although it is well known for its good fairness and welfare properties, it is not strategyproof. In view of this, we address several fundamental questions regarding equilibria under PS. Firstly, we show that Nash deviations under the PS rule can cycle. Despite the possibilities of cycles, we prove that a pure Nash equilibrium is guaranteed to exist under the PS rule. We then show that verifying whether a given profile is a pure Nash equilibrium is coNP-complete, and computing a pure Nash equilibrium is NP-hard. For two agents, we present a linear-time algorithm to compute a pure Nash equilibrium which yields the same assignment as the truthful profile. Finally, we conduct experiments to evaluate the quality of the equilibria that exist under the PS rule, finding that the vast majority of pure Nash equilibria yield social welfare that is at least that of the truthful profile.

1 Introduction

Resource allocation is a fundamental and widely applicable area within AI and computer science. When resource allocation rules are not strategyproof and agents do not have incentive to report their preferences truthfully, it is important to understand the possible manipulations, Nash dynamics, and the existence and computation of equilibria.

In this paper we consider the *probabilistic serial (PS)* rule for the *assignment problem*. In the *assignment problem* we have a possibly unequal number of agents and objects where the agents express preferences over objects and, based on these preferences, the objects are allocated to the agents [Aziz *et al.*, 2014; Bogomolnaia and Moulin, 2001; Gärdenfors, 1973; Hylland and Zeckhauser, 1979]. The model is applicable to many resource allocation and fair division settings where the objects may be public houses, school seats, course enrollments, kidneys for transplant, car park spaces, chores, joint assets, or time slots in schedules. The *probabilistic serial (PS)* rule is a randomized (or fractional) assignment rule. A randomized or fractional assignment rule takes the preferences of the agents into account in order to allocate each agent a fraction of the object. If the objects are indivisible but allocated in a randomized way, the fraction can also be interpreted as the probability of receiving the object. Randomization is widespread in resource allocation as it is a natural way to ensure procedural fairness [Budish *et al.*, 2013].

A prominent randomized assignment rule is the PS rule [Bogomolnaia and Heo, 2012; Bogomolnaia and Moulin, 2001; Budish *et al.*, 2013; Katta and Sethuraman, 2006; Kojima, 2009; Yilmaz, 2010; Saban and Sethuraman, 2014]. PS works as follows: each agent expresses a linear order over the set of houses.[1] Each house is considered to have a divisible probability weight of one. Agents simultaneously and at the same speed eat the probability weight of their most preferred house that has not yet been completely eaten. Once a house has been completely eaten by a subset of the agents, each of these agents starts eating his next most preferred house that has not been completely eaten (i.e., they may "join" other agents already eating a different house or begin eating new houses). The procedure terminates after all the houses have been completely eaten. The random allocation of an agent by PS is the amount of each house he has eaten. Although PS was originally defined for the setting where the number of houses is equal to the number of agents, it can be used without any modification for any number of houses relative to the number agents [Bogomolnaia and Moulin, 2001; Kojima, 2009].

In order to compare random allocations, an agent needs to consider relations between them. We consider two well-known relations between random allocation [Schulman and Vazirani, 2012; Saban and Sethuraman, 2014; Cho, 2012]: (i) *expected utility (EU)*, and (ii) *downward lexicographic (DL)*. For EU, an agent prefers an allocation that yields more expected utility. For DL, an agent prefers an allocation that gives a higher probability to the most preferred alternative that has different probabilities in the two allocations. Throughout the paper, we assume that agents express *strict* preferences over houses, i.e., they are not indifferent between any two houses.

The PS rule fares well in terms of fairness and welfare [Bogomolnaia and Heo, 2012; Bogomolnaia and Moulin, 2001; Budish *et al.*, 2013; Kojima, 2009; Yilmaz, 2010]. It satisfies strong envy-freeness and efficiency with respect to

[1] We use the term house throughout the paper though we stress any object could be allocated with these mechanisms.

the DL relation [Bogomolnaia and Moulin, 2001; Schulman and Vazirani, 2012; Kojima, 2009]. Generalizations of the PS rule have been recommended and applied in many settings [Aziz and Stursberg, 2014; Budish *et al.*, 2013]. The PS rule also satisfies some desirable incentive properties: if the number of houses is at most the number of agents, then PS is DL-strategyproof [Bogomolnaia and Moulin, 2001; Schulman and Vazirani, 2012]. Another well-established rule, *random serial dictator (RSD)*, is not envy-free, not as efficient as PS [Bogomolnaia and Moulin, 2001], and the fractional allocations under RSD are #P-complete to compute [Aziz *et al.*, 2013].

Although PS performs well in terms of fairness and welfare, unlike RSD, it is not strategyproof. Aziz *et al.* [2015] showed that, in the scenario where one agent is strategic, computing his best response (manipulation) under complete information of the other agents' strategies is NP-hard for the EU relation, but polynomial-time computable for the DL relation. In related work, Ekici and Kesten [2012] showed that when agents are not truthful, the outcome of PS may not satisfy desirable properties related to efficiency and envy-freeness. Heo and Manjunath [2012] provided a necessary and sufficient condition for implementability of Nash equilibrium for the random assignment problem. In contrast to the work of Aziz *et al.* [2015], we consider the situation where *all* agents are strategic. We especially focus on pure Nash equilibria (PNE) — reported preferences profiles for which no agent has an incentive to report a different preference. We examine the following natural questions for the first time: *(i) What is the nature of best response dynamics under the PS rule? (ii) Is a (pure) Nash equilibrium always guaranteed to exist? (iii) How efficiently can a (pure) Nash equilibrium be computed? (iv) What is the difference in quality of the various equilibria that are possible under the PS rule?*

Contributions. For the PS rule we show that expected utility best responses can cycle for *any* cardinal utilities consistent with the ordinal preferences. This is significant as Nash dynamics in matching theory has been an active area of research, especially for the stable matching problem [Ackermann *et al.*, 2011], and the presence of a cycle means that following a sequence of best responses is not guaranteed to result in an equilibrium profile. We then prove that a pure Nash equilibrium (PNE) is guaranteed to exist for any number of agents and houses and any utilities. To the best of our knowledge, this is the first proof of the existence of a Nash equilibrium for the PS rule. For the case of two agents we present a linear-time algorithm to compute a preference profile that is in PNE with respect to the original preferences. We show that the general problem for computing a PNE is NP-hard. Finally, we run a set of experiments on real and synthetic preference data to evaluate the welfare achieved by PNE profiles compared to the welfare achieved under the truthful profile.

2 Preliminaries

An *assignment problem* $(N, H, \succ)$ consists of a set of agents $N = \{1, \dots, n\}$, a set of houses $H = \{h_1, \dots, h_m\}$ and a preference profile $\succ = (\succ_1, \dots, \succ_n)$ in which $\succ_i$ denotes a complete, transitive and strict ordering on H representing the preferences of agent i over the houses in H. A *fractional assignment* is an $(n \times m)$ matrix $[p(i)(h_j)]_{1\leq i \leq n, 1 \leq j \leq m}$ such that for all $i \in N$, and $h_j \in H$, $0 \leq p(i)(h_j) \leq 1$; and for all $j \in \{1, \dots, m\}$, $\sum_{i\in N} p(i)(h_j) = 1$. The value $p(i)(h_j)$ is the fraction of house h_j that agent i gets. Each row $p(i) = (p(i)(h_1), \dots, p(i)(h_m))$ represents the *allocation* of agent i. A fractional assignment can also be interpreted as a random assignment where $p(i)(h_j)$ is the probability of agent i getting house h_j.

Given two random assignments p and q, $p(i) \succ_i^{DL} q(i)$ i.e., a player i *DL (downward lexicographic) prefers* allocation $p(i)$ to $q(i)$ if $p(i) \neq q(i)$ and for the most preferred house h such that $p(i)(h) \neq q(i)(h)$, we have that $p(i)(h) > q(i)(h)$. When agents are considered to have cardinal utilities for the houses, we denote by $u_i(h)$ the utility that agent i gets from house h. We will assume that the total utility of an agent equals the sum of the utilities that he gets from each of the houses. Given two random assignments p and q, $p(i) \succ_i^{EU} q(i)$, i.e., a player i *EU (expected utility) prefers* allocation $p(i)$ to $q(i)$ if $\sum_{h\in H} u_i(h) \cdot p(i)(h) > \sum_{h \in H} u_i(h) \cdot q(i)(h)$. Since for all $i \in N$, agent i compares assignment p with assignment q only with respect to his allocations $p(i)$ and $q(i)$, we will sometimes abuse the notation and use $p \succ_i^{EU} q$ for $p(i) \succ_i^{EU} q(i)$.

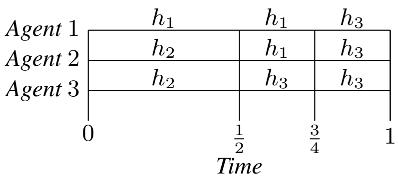
A *random assignment rule* takes as input an assignment problem $(N, H, \succ)$ and returns a random assignment which specifies what fraction or probability of each house is allocated to each agent. We will primarily focus on the expected utility setting but will comment on and use DL wherever needed.

The Probabilistic Serial Rule and Equilibria. The *Probabilistic Serial (PS) rule* is a random assignment algorithm in which we consider each house as infinitely divisible [Bogomolnaia and Moulin, 2001; Kojima, 2009]. At each point in time, each agent is eating (consuming the probability mass of) his most preferred house that has not been completely eaten. Each agent eats at the same unit speed. Hence all the houses are eaten at time m/n and each agent receives a total of m/n units of houses. The probability of house h_j being allocated to i is the fraction of house h_j that i has eaten. The PS fractional assignment can be computed in time $O(mn)$. The following example from Bogomolnaia and Moulin; Aziz *et al.* [2001; 2015] shows how PS works.

Example 1 (PS rule). *Consider an assignment problem with the following preference profile.*

$$\succ_1: \quad h_1, h_2, h_3$$
$$\succ_2: \quad h_2, h_1, h_3$$
$$\succ_3: \quad h_2, h_3, h_1$$

Agents 2 *and* 3 *start eating* h_2 *simultaneously while agent* 1 *eats* h_1*. When* 2 *and* 3 *finish* h_2 *at time* $1/2$*, each having consumed* $1/2$ *of* h_1*, agent* 3 *has only eaten half of* h_1*. Since agent 2 prefers* h_1 *to* h_3 *and* h_1 *has not been completely consumed, agent 2 joints agent 1 in consuming the remaining part of* h_1 *while agent 3 begins to eat* h_3*. Agent 1 and 2 finish consuming* *the remaining* $1/2$ *of* h_1 *at time* $3/4$*; having consumed an additional quarter of* h_1 *each. Since all the houses are completely eaten except* h_3*, agents 1 and 2 join agent 3 in finishing* h_3*. The timing of the eating can be seen below.*



The final allocation computed by PS is

$$PS(\succ_1, \succ_2, \succ_3) = \begin{pmatrix} 3/4 & 0 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{pmatrix}.$$

Consider the assignment problem in Example 1. If agent 1 misreports his preferences as follows: $\succ_1'$: h_2, h_1, h_3, then

$$PS(\succ_1', \succ_2, \succ_3) = \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \\ 0 & 1/3 & 2/3 \end{pmatrix}.$$

If we suppose that $u_1(h_1) = 7$, $u_1(h_2) = 6$, and $u_1(h_3) = 0$, then agent 1 gets more expected utility when he reports $\succ_1'$. In the example, the truthful profile is in PNE with respect to DL preferences but not expected utility.

We study the existence and computation of Nash equilibria. For a preference profile $\succ$, we denote by $(\succ_{-i}, \succ_i')$ the preference profile obtained from $\succ$ by replacing agent i's preference by $\succ_i'$.

3 Nash Dynamics

When considering Nash equilibria of any setting, one of the most natural ways of proving that a PNE always exists is to show that better or best responses do not cycle which implies that eventually, Nash dynamics terminate at a Nash equilibrium profile. Our first result is that DL and EU best responses can cycle. For EU best responses, this is even the case when agents have Borda utilities.

Theorem 1. *With 2 agents and 5 houses where agents have Borda utilities, EU best responses can lead to a cycle in the profile.*

Proof. The following 5 step sequence of best responses leads to a cycle. We use U to denote the matrix of utilities of the agents over the houses such that $U[1, 1]$ is the utility of agent 1 for house h_1. Note that P starts as the truthful reporting in our example. The initial state of the agents is:

$$\begin{matrix} \succ_1: & h_2, h_3, h_5, h_4, h_1 \\ \succ_2: & h_5, h_3, h_4, h_1, h_2 \end{matrix} \quad U_0 = \begin{pmatrix} 0 & 4 & 3 & 1 & 2 \\ 1 & 0 & 3 & 2 & 4 \end{pmatrix}.$$

This yields the following allocation and utilities at the start:

$$PS(\succ_1, \succ_2) = \begin{pmatrix} 1/2 & 1 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 1/2 & 1 \end{pmatrix}, EU_0 = \begin{pmatrix} 6 \\ 7 \end{pmatrix}.$$

(1.) Agent 1 deviates to increase his utility. He reports the preference $\succ_1'$: h_3, h_4, h_2, h_1, h_5; which results in

$$PS(\succ_1', \succ_2) = \begin{pmatrix} 0 & 1 & 1 & 1/2 & 0 \\ 1 & 0 & 0 & 1/2 & 1 \end{pmatrix}, EU_1 = \begin{pmatrix} 7.5 \\ 6 \end{pmatrix}.$$

(2.) Agent 2 changes his report to $\succ_2'$: h_3, h_4, h_5, h_1, h_2. This increases the utility of agent 2 to 7 and decreases the utility of agent 1 to 6.

(3.) Agent 1 changes his report to $\succ_1''$: h_3, h_5, h_2, h_1, h_4. This increases the utility of agent 1 to 7.5 and decreases the utility of agent 2 to 4.5.

(4.) Agent 2 changes his report to $\succ_2''$: h_5, h_3, h_4, h_1, h_2. which increases his expected utility to 6.5 while decreasing the expected utility of agent 1 to 7.

(5.) Agent 1 changes his report to $\succ_1'''$: h_3, h_4, h_2, h_1, h_5. Notice that $\succ_1''' = \succ_1'$ and $\succ_2'' = \succ_2$. This is the same profile as the one of Step 1, so we have cycled. □

It can be verified that every response in the example in the proof above is also a DL best response. Since for the case of two agents and PS, a DL best response is equivalent to an EU best response for any cardinal utilities consistent with the ordinal preferences [Aziz *et al.*, 2015], it follows that DL best responses and EU best responses (with respect to any cardinal utilities consistent with the ordinal preferences) can cycle.

The fact that best responses can cycle means that simply following best responses need not result in a PNE. Hence the normal form game induced by the PS rule is not a potential game [Monderer and Shapley, 1996]. Checking whether an instance has a Nash equilibrium appears to be a challenging problem. The naive method requires going through $O(m!^n)$ profiles, which is super-polynomial even when $n = O(1)$ or $m = O(1)$.

4 Existence of Pure Nash Equilibria

Although it seems that computing a Nash equilibrium is a challenging problem (we give hardness results in the next section), we show that at least one pure Nash equilibrium is guaranteed to exist for any number of houses, any number of agents, and any preference relation over fractional allocations.[2] The proof relies on showing that the PS rule can be modelled as a perfect information extensive form game.

Theorem 2. *A PNE is guaranteed to exist under the PS rule for any number of agents and houses, and for any relation between allocations.*

Proof. Consider running PS on all possible $m!^n$ preference profiles for n agents and m houses. In each profile i, let $t_i^1, \ldots, t_i^{k_i}$ be the k_i different time points in the PS algorithm run for the i-th profile when at least one house is finished. Note that all such time points are rationals: t_i^1 is rational and if $t_i^1, \ldots, t_i^j$ are rational, then t_i^{j+1} is rational as well. For non-zero rationals $r_1, \ldots, r_n$, the value $\text{GCD}(r_1, \ldots, r_n)$ denotes the greatest rational number r for which all the r_i/r are integers. Let $g = \text{GCD}(\{t_i^{j+1} - t_i^j : j \in \{1, \ldots, k_i - 1\}, i \in \{1, \ldots, m!^n\})$. Since in each profile i, $t_i^{j+1} - t_i^j > 0$ and rational for all $j \in \{0, \ldots, k_i - 1\}$, g is well-defined.

The time interval length g is small enough such that each run of the PS rule can be considered to be a 'discrete' rule with m/g stages of duration g. Each stage can be viewed as

[2]We already know from Nash's original result that a *mixed* Nash equilibrium exists for any game.

having n sub-stages so that in each stage, agent i eats g/n units of a house in sub-stage i of a stage. In each sub-stage only one agent eats g/n units of the most favoured house that is available. Hence we now view PS as consisting of a total of mn/g sub-stages and the agents keep coming in order $1, 2, \ldots, n$ to eat g units of the most preferred house that is still available. If an agent eats g units of a house in a stage then it will eat g units of the same house in his sub-stage of the next stage as long as the house has not been fully eaten. Therefore, the discretized version of PS is equivalent to the original PS because g is small enough. Hence if there is preference profile that is a PNE in the discretized version of the PS, it is also a PNE for the original PS rule.

Consider a perfect information extensive form game tree. For a fixed reported preference profile, the PS rule unravels accordingly along a unique path starting at the root and ending at a leaf. Each level of the tree represents a sub-stage in which a certain agent has his turn to eat g/n units of his most preferred available house. Note that there is a one-to-one correspondence between the paths in the tree and the ways the PS algorithm unravels, depending on the reported preferences. For each path in the tree, we know the house eaten by each agent at each time point. For the other direction, for each unique way the agents eat the houses during the PS algorithm, there is a unique path in the extensive game tree.

A subgame perfect Nash equilibrium (SPNE) is guaranteed to exist for the discretized version of PS via backward induction: starting from the leaves and moving towards the root of the tree, the agent at the specific node chooses an action that maximizes his utility given the actions determined for the children of the node. The SNPE identifies at least one such path from a leaf to the root of the game. The path can be used to read out the most preferred house of each agent at each point. The preference of each agent i consistent with the path is then constructed by listing houses in the order in which i eats them, starting from the root of the tree to the leaf. Those houses that an agent did not eat at all can be placed arbitrarily at the end of the preference list. Such a preference profile is in pure Nash equilibrium under the discretized version of PS because it is an SPNE. Hence, such a preference profile is also a PNE for the actual PS rule. □

In Figure 1, we illustrate the perfect information extensive form game corresponding to PS.

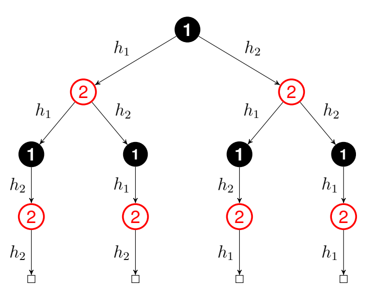


Figure 1: A perfect information extensive form game based on breaking down the PS algorithm into a series of stages. For two agents, each agent eats half a house in his turn.

5 Complexity of Pure Nash Equilibrium

Our argument for the existence of a Nash equilibrium is constructive. However, naively constructing the extensive form game and then computing a subgame perfect Nash equilibrium requires exponential space and time. It is unclear whether a sub-game perfect Nash equilibrium or any Nash equilibrium preference profile can be computed in polynomial time.

5.1 General Complexity Results

In this section, we show that computing a PNE is NP-hard and verifying whether a profile is a PNE is coNP-complete. Recently it was shown that computing an expected utility best response is NP-hard [Aziz *et al.*, 2015]. Since equilibria and best responses are somewhat similar, one would expect that problems related to equilibria under PS are also hard. However, there is no simple reduction from best response to equilibria computation or verification. In view of this, we prove results regarding PNE by closely analyzing the reduction given by Aziz *et al.* [2015]. First, we show that checking whether a given preference profile is in PNE under the PS rule is coNP-complete.

Theorem 3. *Given agents' utilities, checking whether a given preference profile is in PNE under the PS rule is coNP-complete.*

Proof. The problem is in coNP, since a Nash deviation is a polynomial time checkable No-certificate.

We leverage the reduction given by Aziz *et al.* [2015] which shows that computing an expected utility best response for a single agent is NP-hard for the PS rule. Their reduction considers one manipulator (agent 1) while the other agents $N \setminus \{1\}$ are 'non-manipulators'. We show that checking whether the truthful preference profile is in PNE is coNP-complete.

The reduction given by Aziz *et al.* [2015] is from a variant of 3SAT where each literal appears twice. They show that given an assignment setting and a utility function for agent 1 it is NP-hard to determine if agent 1 can report preferences that reach a target utility T. The reduction creates an instance of PS that is best conceptualized as having one main part, where the manipulator (agent 1) is present, and 17 duplicate parts, where the dummy manipulators are present. Each of these duplicate parts can be further broken down into a set of choice rounds, corresponding to the number of variables in the 3SAT instance, and a final clause round. Houses associated to a round will be eaten before progressing to the next round. At first all agents are synchronised, that is no agent starts eating a house from the next round whilst others are still eating from houses from this round. As the algorithm progresses, some agents end up being slightly desynchronised, which means they will start eating houses from the next round some small amount of time before all houses from the previous rounds have been eaten. The set of agents is comprised of the manipulator, the set of dummy manipulators, and a pair of agents for each literal in the 3SAT formula for each duplicate part. The set of houses is comprised of a set of slowdown houses, shared between all the duplicate parts and used to synchronize all the agents; a prize house that the manipulator wants to eat as early as possible after consuming the choice houses, a set of consolation prize houses for each one of the dummy manipulators, a house for each literal in the 3SAT instance for each round for each of the parts; and three houses for each clause in the 3SAT formula.

Intuitively, during each choice round the manipulator must choose between the positive and negative literal by eating the corresponding house. To ensure that all the different parts of the reduction stay synchronized the slowdown houses are used between rounds. The utility of the prize house is set in a way such that the manipulator can reach a target utility T or more if and only if he selects a satisfying assignment for the all the clauses of the 3SAT formula. Any selection of literals by agent 1 that does not satisfy the formula will leave a set of agents free to consume a large portion of the prize house before agent 1 can reach it. Hence, finding an expected utility best response for PS is NP-hard.

In the original reduction, the utility functions of agents in $N' = N \setminus \{1\}$ are not specified. To prove that checking if the truthful preference profile is in PNE under the PS rule we specify the utility function of agents in N' as follows: the utility of an agent in N' for his j-th most preferred house is $(4cn)^{m-j+1}$, where $n = |N|$, m is the number of houses and c is an integer constant greater than 1000. These utility functions can be represented in space that is polynomial in $O(n+m)$. We rely on two main observations about the original reduction.

Observation 1. *In the truthful profile, whenever an agent finishes eating a house, all houses have either been fully allocated or they are only at most $\frac{c-1}{c}$ eaten, where $c > 1000$ is an integer constant.*

This observation relies on the fact that the constructed PS instance consists of several rounds. No matter how big the instance is the desynchronisation between the agents in those rounds is given by a constant number of values. This allows us to bound the smallest fraction of a house that is left at any point in a round where an agent finishes eating a house by at least $1/c$.

Observation 2. *In the truthful profile every house except the prize house (the last house that is eaten) is eaten by at least 2 agents.*

Again this can be seen simply by noting that during the choice rounds each agent is paired with at least one other agent. During the final round, all literals corresponding to a clause will eat the house associated with that clause. The agents cannot be desynchronised to a point where one literal agent eats a whole clause house.

We now show that due to the utility function constructed, each agent from N' is compelled to report truthfully. Assume for contradiction that this is not the case, and let us consider the earliest house (when running the PS rule) that some agent $i \in N'$ starts to eat although he prefers another available house h. Let k denote the number of agents who eat a fraction of h under the truthful profile. By reporting truthfully, we show that agent i can get $\frac{1/n-1/2n}{c} = 1/2cn$ more of h than by delaying eating h. Let us consider how much additional fraction of h agent i can consume by reporting truthfully. If he reports truthfully, he can start eating h earlier and, in the worst case, he can only start $1/cn$ time units earlier by Observation 1. This means that h is consumed earlier by a time of $1/cn$ if i reports truthfully. Consider the time interval of length $1/cn$ between the time when h is finished when i is truthful about h and the time h is finished when i delays eating h. In this last stretch of time interval $1/cn$, i gets $\frac{1}{k} \cdot \frac{1}{cn}$ of h extra when he does not report truthfully. Hence by reporting truthfully, i gets at least $\frac{1/n-1/kn}{c}$ more of h which is at least $1/2cn$ since $k \geq 2$ by Observation 2. Due to the utilities constructed, even if i gets all the less preferred houses, he cannot make up for the loss in utility for getting only $1/2cn$ of h.

We have established above that the agents in N' report truthfully in *each* PNE. This implies that the truthful preference profile is in PNE iff agents in $N' = N \setminus \{1\}$ report truthfully and agent 1's truthful report is his best EU response. Assuming that the agents in $N \setminus \{1\}$ report truthfully, checking whether the truthful preference is agent 1's best response was shown to be NP-hard. We have already shown that the agents N' report truthfully in a PNE. Hence checking whether the truthful profile is in PNE is coNP-hard. □

Next, we show that computing a PNE with respect to the underlying utilities of the agents is NP-hard.

Theorem 4. *Given agent's utilities, computing a preference profile that is in PNE under the PS rule is NP-hard.*

Proof. The same argument as above shows that the agents in N' report truthfully in a PNE. Hence, a preference profile is in PNE iff agent 1 reports his EU best response and the other agents report truthfully. It has already been shown that computing this EU best response is NP-hard [Aziz *et al.*, 2015] when the other agents are $N \setminus \{1\}$ and report truthfully. Thus computing a PNE is NP-hard. □

5.2 Case of Two Agents

In this section, we consider the interesting and common special case of just two agents. Since an EU best response can be computed in linear time for the case of two agents [Aziz *et al.*, 2015], it follows that it can be verified whether a profile is a PNE in polynomial time as well.

We can prove the following theorem for the "threat profile" whose construction is shown in Algorithm 1. The idea is that by placing houses in the appropriate place in the other agent's preference list, we ensure that the agent feels threatened that his most preferred house will be eaten by the other agent so that he eats his most preferred available house first.

Theorem 5. *Under PS and for two agents, there exists a preference profile that is in DL-Nash equilibrium and results in the same assignment as the assignment based on the truthful preferences. Moreover, it can be computed in linear time.*

Proof. The proof is by induction over the length of the constructed preference lists. The main idea of the proof is that if both agents compete for the same house then they do not have an incentive to delay eating it. If the most preferred houses do not coincide, then both agents get them with probability one

Input: $(\{1,2\}, H, (\succ_1, \succ_2))$
Output: The *"threat profile"* (Q_1, Q_2) where Q_i is the preference list of agent i for $i \in \{1,2\}$.

Let P_i be the preference list of agent $i \in \{1,2\}$ with first(P_i) being the most preferred house in P_i for agent i.
Initialise Q_1 and Q_2 to empty lists.
while P_1 and P_2 are not empty **do**
    Let $h = \text{first}(P_1)$ and $h' = \text{first}(P_2)$
    Append h to the end of Q_1; Append h' to the end of Q_2
    Delete h and h' from P_1 and P_2
    if $h \neq h'$ **then**
        Append h' to the end of Q_1; Append h to the end of Q_2;
return (Q_1, Q_2).

Algorithm 1: Threat profile DL-Nash equilibrium for 2 agents (which also is an EU-Nash equilibrium) which provides the same allocation as the truthful profile.

but will not get them completely if they delay eating them. The algorithm is described as Algorithm 1.

We now prove that Q_1 is a DL best response against Q_2 and Q_2 is a DL best response against Q_1. The proof is by induction over the length of the preference lists. For the first elements in the preference lists Q_1 and Q_2, if the elements coincide, then no agent has an incentive to put the element later in the list since the element is both agents' most preferred house. If the maximal elements do not coincide i.e. $h \neq h'$, then 1 and 2 get h and h' respectively with probability one. However they still need to express these houses as their most preferred houses because if they don't, they will not get the house with probability one. The reason is that h is the next most preferred house after h' for agent 2 and h' is the next most preferred house after h for agent 1. Agent 1 has no incentive to change the position of h' since h' is taken by agent 2 completely before agent 1 can eat it. Similarly, agent 2 has no incentive to change the position of h since h is taken by agent 1 completely before agent 2 can eat it. Now that the positions of h and h' have been completely fixed, we do not need to consider them and can use induction over Q_1 and Q_2 where h and h' are deleted. □

The desirable aspect of the threat profile is that since it results in the same assignment as the assignment based on the truthful preferences, the resulting assignment satisfies all the desirable properties of the PS outcome with respect to the original preferences. Since a DL best response algorithm is also an EU best response algorithm for the case of two agents [Aziz *et al.*, 2015], we get the following corollary.

Corollary 1. *Under PS and for 2 agents, there exists a preference profile that is in Nash equilibrium for any utilities consistent with the ordinal preferences. Moreover it can be computed in linear time.*

6 Experiments

We conducted a series of experiments to understand the number and quality of equilibria that are possible under the PS rule. For quality, we use the utilitarian social welfare (SW) function, i.e., the sum of agent utilities. We are limited by the large search space needed to examine equilibria. For instance, for each set of cardinal preferences we generate, we consider all misreports ($m!$) for all agents (n) leaving us with a search space of size $m!^n$ for each of the samples for each combination of parameters. Thus, we only report results for small numbers of agents and houses in this section. We generated 1000 samples for each combination of preference model, number of agents, and number of items; reporting the aggregate statistics for these experiments for only the 4 agent case in Figures 2 and 3; the results for $n = 2$ and $m \in \{2, \ldots, 5\}$ as well as $n = 3$ and $m \in \{2, 3, 4\}$ are similar. Each individual sample with 4 agents and 4 houses took about 15 minutes to complete using one core on an Intel Xeon E5405 CPU running at 2.0 GHz with 4 GB of RAM running Debian 6.0 (build 2.6.32-5-amd64 Squeeze10). The total compute time for 1000 samples for each of 6 models was over 40 days.

We used a variety of common statistical models to generate data (see, e.g., [Mattei, 2011; Mallows, 1957; Lu and Boutilier, 2011; Berg, 1985]): the Impartial Culture (**IC**) model generates all preferences uniformly at random; the Single Peaked Impartial Culture (**SP-IC**) generates all preference profiles that are single peaked uniformly at random [Walsh, 2015]; Mallows Models (**Mallows**) is a correlated preference model where the population is distributed around a reference ranking proportional to the Kendall-Tau distance; Polya-Eggenberger Urn Models (**Urn**) creates correlations between the agents, once a preference order has been randomly selected, it is subsequently selected with higher probability. In our experiments we set the probability that the second order is equivalent to the first to 0.5. For generating all model data we used the PrefLib Tool Suite [Mattei and Walsh, 2013]. We also used real world data from PrefLib [Mattei and Walsh, 2013]: AGH Course Selection (ED-00009). This data consists of students bidding on courses to attend in the next semester. We sampled students from this data (with replacement) as the agents after we restricted the preference profiles to a random set of houses of a specified size.

To compare the different allocations achieved under PS we need to give each agent not only a preference order but also a utility for each house. Formally we have, for all $i \in N$ and all $h_j \in H$, a value $u_i(h_j) \in \mathbb{R}$. To generate these utilities we use what we call the *Random* model: we uniformly at random generate a real number between 0 and 1 for each house. We sort this list in strictly decreasing order, if we cannot, we generate a new list (every sample we generated was a strict order in our testing). We normalize these utilities such that each agent's utility sums to a constant value (here, the number of houses) that is the same for all agents. We found the Random utility model to be the most manipulable and admit the worst equilibria. Therefore, we only focus on this utility model here (over Borda or Exponential utilities) as it represents, empirically, a worst case. We separate equilibria into three categories: those where the SW is the same as in the truthful profile, those where we have a decrease in SW, and those where we have an increase in SW. Given the social welfare of two different profiles, SW_1 and SW_2, we use percentage change ($\frac{|SW_1 - SW_2|}{SW_1} \cdot 100$) to understand the magnitude of this difference.

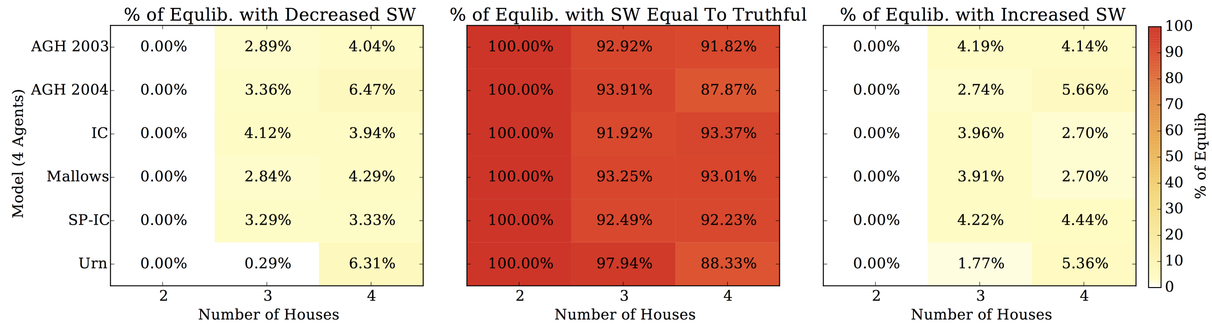
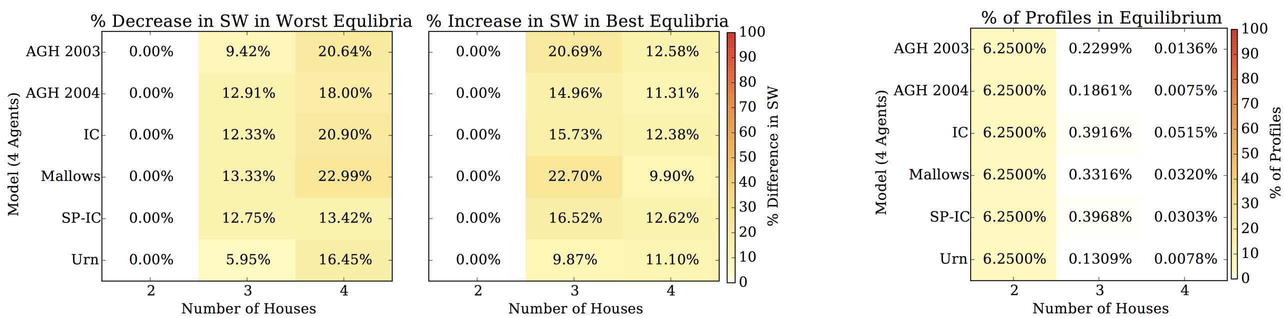


Figure 2: Classification of equilibria for all 1000 samples per setting with four agents ($n = 4$), 2 to 4 houses ($m \in \{2, 3, 4\}$), and preferences drawn from the six models. We can see that the vast majority of the equilibria found across all samples have the same social welfare as the truthful profile. In general, there are roughly the same number of equilibria that increase as those that decrease it.



(A) (B)

Figure 3: (A) The maximum and minimum percentage increase or decrease in social welfare over all 1000 samples in settings with four agents ($n = 4$), 2 to 4 houses ($m \in \{2, 3, 4\}$), and preferences drawn from the six models. We see that for three houses the gain of the best profile is, in general, slightly more than the loss in the worst profile with respect to the truthful profile; this trend appears to reverse for settings with four houses. (B) The average number of the $m!^4$ profiles that are in equilibria per sample with four agents ($n = 4$), 2 to 4 houses ($m \in \{2, 3, 4\}$). The more uncorrelated models (i.e., IC and SP-IC) admit the highest number of equilibria.

For all models, for all combinations of $n \in \{2, 3, 4\}$ agents and $m \in \{2, 3, 4\}$ houses there are, generally, slightly more equilibria that increase social welfare compared to the truthful profile than those that decrease it, as illustrated in Figure 2 (four agents only). However, the vast majority of equilibria have the same social welfare as the truthful profile, and the best and worst equilibria change the SW up or down roughly the same magnitude, as illustrated in Figure 3. Hence, if any or all of the agents manipulate, there may be a loss of SW at equilibria, but there is also the potential for gains; and the most common outcome of all these agents being strategic is that, dynamically, we will wind up in an equilibria which provides the same SW as the truthful one. Our main observations are: (i) The vast majority of equilibria have social welfare equal to the social welfare in the truthful profile. (ii) In general, the number of PNE that have increased social welfare (with respect to the truthful profile) is slightly more than the number of PNE that have decreased social welfare. (iii) The maximum increase and decrease in SW in equilibria compared to the truthful profile was observed to be less than 23% either way. (iv) There are very few profiles that are in equilibria, overall. Profiles with relatively high degrees of correlation between the preferences (Urn and AGH 2004) have fewer equilibrium profiles than the less correlated models (IC and SP-IC). (v) These trends appear stable with small numbers of agents and houses. We observed similar results for all combinations of $n \in \{2, 3, 4\}$ agents and $m \in \{2, 3, 4\}$ houses.

7 Conclusions

We conducted a detailed analysis of strategic aspects of the PS rule including the complexity of computing and verifying PNE. The fact that PNE are computationally hard to compute in general may act as a disincentive or barrier to strategic behavior. Our experimental results show PS is relatively robust, in terms of social welfare, even in the presence of strategic behaviour. Our study leads to a number of new research directions. It will be interesting to extend our algorithmic results to the extension of PS for indifferences [Katta and Sethuraman, 2006]. Studying *strong* Nash equilibria and a deeper analysis of Nash dynamics are other interesting directions.

Acknowledgments

NICTA is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program. Serge Gaspers is the recipient of an Australian Research Council Discovery Early Career Researcher Award (project number DE120101761) and a Future Fellowship (project number FT140100048).

References

[Ackermann *et al.*, 2011] H. Ackermann, P. W. Goldberg, V. S. Mirrokni, H. Röglin, and B. Vöcking. Uncoordinated two-sided matching markets. *SIAM Journal on Computing*, 40(1):92–106, 2011.

[Aziz and Stursberg, 2014] H. Aziz and P. Stursberg. A generalization of probabilistic serial to randomized social choice. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI)*, pages 559–565. AAAI Press, 2014.

[Aziz *et al.*, 2013] H. Aziz, F. Brandt, and M. Brill. The computational complexity of random serial dictatorship. *Economics Letters*, 121(3):341–345, 2013.

[Aziz *et al.*, 2014] H. Aziz, S. Gaspers, S. Mackenzie, and T. Walsh. Fair assignment of indivisible objects under ordinal preferences. In *Proceedings of the 13th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 1305–1312, 2014.

[Aziz *et al.*, 2015] H. Aziz, S. Gaspers, S. Mackenzie, N. Mattei, N. Narodytska, and T. Walsh. Manipulating the probabilistic serial rule. In *Proceedings of the 14th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 1451–1459, 2015.

[Berg, 1985] S. Berg. Paradox of voting under an urn model: The effect of homogeneity. *Public Choice*, 47:377–387, 1985.

[Bogomolnaia and Heo, 2012] A. Bogomolnaia and E. J. Heo. Probabilistic assignment of objects: Characterizing the serial rule. *Journal of Economic Theory*, 147:2072–2082, 2012.

[Bogomolnaia and Moulin, 2001] A. Bogomolnaia and H. Moulin. A new solution to the random assignment problem. *Journal of Economic Theory*, 100(2):295–328, 2001.

[Budish *et al.*, 2013] E. Budish, Y.-K. Che, F. Kojima, and P. Milgrom. Designing random allocation mechanisms: Theory and applications. *American Economic Review*, 103(2):585–623, 2013.

[Cho, 2012] W. J. Cho. Probabilistic assignment: A two-fold axiomatic approach. Mimeo, 2012.

[Ekici and Kesten, 2012] O. Ekici and O. Kesten. An equilibrium analysis of the probabilistic serial mechanism. Technical report, Özyeğin University, Istanbul, May 2012.

[Gärdenfors, 1973] P. Gärdenfors. Assignment problem based on ordinal preferences. *Management Science*, 20:331–340, 1973.

[Heo and Manjunath, 2012] E. J. Heo and V. Manjunath. Probabilistic assignment: Implementation in stochastic dominance nash equilibria. Technical Report 1809204, SSRN, 2012.

[Hylland and Zeckhauser, 1979] A. Hylland and R. Zeckhauser. The efficient allocation of individuals to positions. *The Journal of Political Economy*, 87(2):293–314, 1979.

[Katta and Sethuraman, 2006] A-K. Katta and J. Sethuraman. A solution to the random assignment problem on the full preference domain. *Journal of Economic Theory*, 131(1):231–250, 2006.

[Kojima, 2009] F. Kojima. Random assignment of multiple indivisible objects. *Mathematical Social Sciences*, 57(1):134—142, 2009.

[Lu and Boutilier, 2011] T. Lu and C. Boutilier. Learning Mallows models with pairwise preferences. In *Proceedings of the 28th International Conference on Machine Learning (ICML)*, pages 145–152, 2011.

[Mallows, 1957] C. L. Mallows. Non-null ranking models. *Biometrika*, 44(1/2):114–130, 1957.

[Mattei and Walsh, 2013] N. Mattei and T. Walsh. PrefLib: A library for preference data. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT)*, volume 8176 of *Lecture Notes in Artificial Intelligence (LNAI)*, pages 259–270. Springer, 2013.

[Mattei, 2011] N. Mattei. Empirical evaluation of voting rules with strictly ordered preference data. In *Proceedings of the 2nd International Conference on Algorithmic Decision Theory (ADT)*, pages 165–177. 2011.

[Monderer and Shapley, 1996] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14(1):124–143, 1996.

[Saban and Sethuraman, 2014] D. Saban and J. Sethuraman. A note on object allocation under lexicographic preferences. *Journal of Mathematical Economics*, 50:283–289, 2014.

[Schulman and Vazirani, 2012] L. J. Schulman and V. V. Vazirani. Allocation of divisible goods under lexicographic preferences. Technical Report arXiv:1206.4366, arXiv.org, 2012.

[Walsh, 2015] T. Walsh. Generating single peaked votes. *CoRR*, abs/1503.02766, 2015.

[Yilmaz, 2010] O. Yilmaz. The probabilistic serial mechanism with private endowments. *Games and Economic Behavior*, 69(2):475–491, 2010.