Noncooperative Coalition Formation and a Coalition Formation Mechanism

by

Eugene Vorobeychik, Myrna Wooders, Greg Leo

Matthew Chambers, Chen Hajaj, Jian Lou, Martin van der Linden,

for Presentation at the

31st International Conference of the Game Theory Society

Stony Brook, New York

For any questions or comments, including related literature, or for a copy of the paper,

please send me an email at myrna.wooders@gmail.com

We model decentralized coalition formation as a sequential bargaining game, with preferences over outcomes.

Main (Surprising!) Result: If each proposer can make enough (but a finite number) proposals then the SPNE is individually rational and Pareto optimal.

Related papers:

Many (most) make a consequential assumption about the order of proposers – the first player to reject a proposal becomes the next proposer.

Many (most) start with a TU game and develop a bargaining procedure to obtain outcomes in the core.

Coalition Formation as in Perry and Reny (1994), Seidman and Winter (1998), Ray and Vohra (1999), Chatterjee et al. (1993), Okada (1996), Block and Diamantoudi (2006).

The Components of our Accept Reject Game

$N = \{1, \ldots, n\}$: a set of players

$T \in 2^n$: the set of admissible coalitions.

Note: an admissible coalition for i must contain i.

π : a partition of the total player set into coalitions.

π_i: the coalition containing player i

$\succ_i$: the preference of i over i's admissible set of coalitions.

$\succ = (\succ_1, ..., \succ_i, .. \succ_i)$: a preference profile.

Non-Cooperative Coalition Formation Game

An *Accept-Reject-Game* (an ARG) is a game *of perfect information in extensive form with the above components and an ordering* $O = (o_1, o_2, ..., o_n)$ *over the players including each* $i \in N$ *at least once;*

O determines the order in which players can make proposals to other players (or to themselves alone) to form a coalition.

The Game:

o_1 makes a proposal to the (other) members of some subset of players, say S.

The members of S sequentially accept A or reject R the proposal.

It accepted by all members of S, then S is formed and its members have no further actions.

If one member of S rejects the proposal, then the next player in the ordering (who may be the same proposer) proposes.

For the special case of 2-player coalitions: $|N| = 3$.

Suppose each player can only make one proposal.

Suppose i proposes the coalition $\{i, j\}$.

Now suppose i can make only one proposal and proposes $\{i, j\}$.

If j rejects, then it is the turn of the next player in O, say ℓ, to make a proposal– to any subset of players who are not yet in coalitions, i and k.

If 3 accepts then the coalition $\{j, k\}$ is formed.

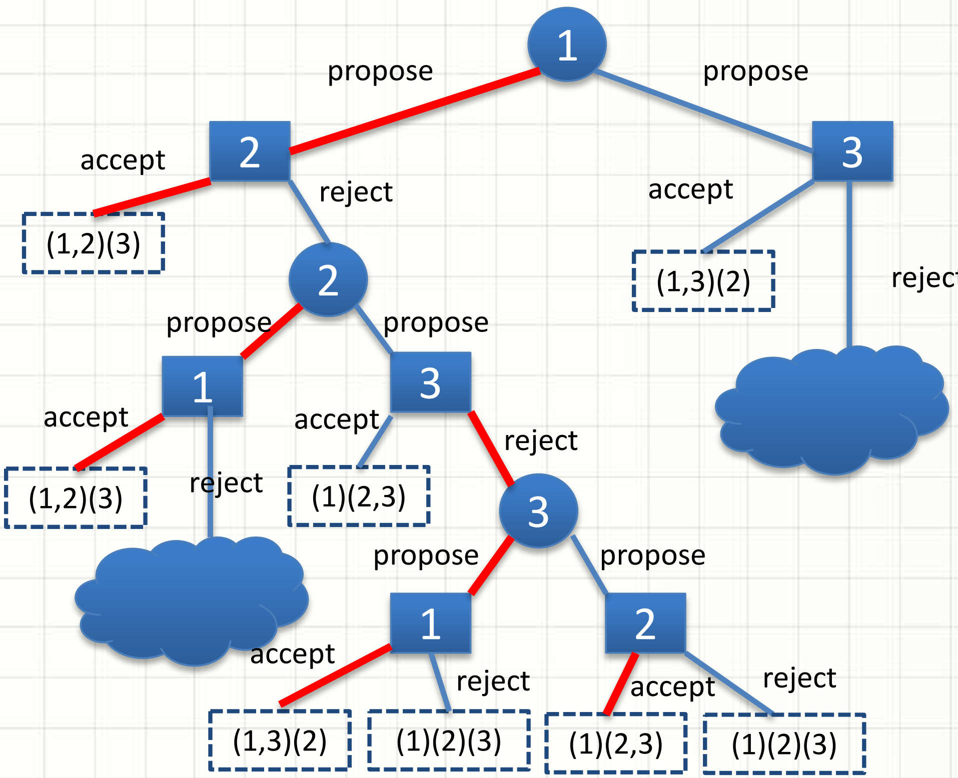
If 3 rejects, then the outcome consists of singleton coalitions $\{i\}$, $\{j\}$, $\{k\}$.

The next diagram illustrates an ARG and its SPNE. The clouds indicate where it is left to the reader to complete the diaagram. It is also used to illustrate that a SPNE is not Pareto Optimal and the slide after that discusses how, by allowing player 1 to make 2 proposals the resulting SPNE is Pareto Optimal.

Computing subgame perfect equilibrium

Proposing order: O= (1, 2, 3)

$$1: \{1,2\} \succ_1 \{1,3\} \succ_1 \{1\}$$
$$2: \{2,3\} \succ_2 \{2,1\} \succ_2 \{2\}$$
$$3: \{3,1\} \succ_3 \{3,2\} \succ_3 \{3\}$$



Equilibrium Properties of ARGs

Is is easy to see that individual rationality is satisfied by the equilibrium of an ARG.,

Assume that players prefer to form a coalition earlier rather than later in the game.

Theorem 1. The SPNE outcome is unique and in an SPNE strategy profile all proposals are accepted.

Surprisingly, Pareto optimally is not necessarily satisfied by an SPNE outcome.

Compare with serial dictatorship (usually attributed to Satherthwaite and to several papers due to Bogomolnania & Moulin), which satisfies Pareto optimality.

Example: Consider a roommate problem with a set of 6 players, $\{1, \ldots, 6\}$ who have the following preferences:

$$
\begin{array}{llllllllll}
1: & \{1,3\} & \succ_1 & \{1,4\} & \succ_1 & \{1,5\} & \succ_1 & \{1,2\} & \succ_1 & \{1,6\} \\
2: & \{2,1\} & \succ_2 & \{2,5\} & \succ_2 & \{2,4\} & \succ_2 & \{2,3\} & \succ_2 & \{2,6\} \\
3: & \{3,2\} & \succ_3 & \{3,4\} & \succ_3 & \{3,1\} & \succ_3 & \{3,5\} & \succ_3 & \{3,6\} \\
4: & \{4,3\} & \succ_4 & \{4,1\} & \succ_4 & \{4,5\} & \succ_4 & \{4,2\} & \succ_4 & \{4,6\} \\
5: & \{5,2\} & \succ_5 & \{5,4\} & \succ_5 & \{5,1\} & \succ_5 & \{5,3\} & \succ_5 & \{5,6\} \\
6: & \{6,5\} & \succ_6 & \{6,1\} & \succ_6 & \{6,2\} & \succ_6 & \{6,3\} & \succ_6 & \{6,4\}
\end{array}
$$

The non-Pareto-optimal SPNE outcome: $\{(1,5),(2,4),(3,6)\}$.

$\{(2,5),(1,4),(3,6)\}$ is a Pareto improvement.

$\{(1,3),(2,5),(4,6)\}$, is Pareto-optimal.

Intuition: The outcome if 1's first proposal were rejected is a credible threat to player 3..

A change in the ordering yields a Pareto-optimal outcome.

Example: Let player 1 make 2 proposals.

If 1's 1st proposal is rejected, the subgame is identical to the game above.

Suppose that 1 first proposes to 3.

If 3 rejects, 3 is coalitioned with 6 in the resulting subgame.

3 strictly prefers $(1, 3)$ to $(3, 6)$ and so would accept 1's first proposal.

Given $(1, 3)$, 2 and 5 are soulmates so $(2, 5)$ would form.

The SPNE outcome, $\{(1, 3), (2, 5), (4, 6)\}$, is Pareto - optimal.

The Coalitional Game

The players N and their preferences $\succ$ determine a co-operative game.

The terminal outcomes yield the payoffs to coalitions.

A coalition structure π is in the *core* if there does not exist a coalition of players $T \subset N$ such that for all $i \in T$, $T \succ_i \pi_i$.

The core may be empty.

Iterated matching of soulmates (IMS) captures the idea that a set of players who, among the set of players not already in coalitions, all prefer to be with each other, are naturally matched.

An example of IMS for a 2-player roommate problem:

Alice and Alex prefer to be with each other rather than anyone else.

They are (1st order) *soulmates* and paired by IMS.

Bertie and Ben are not so smitten. Ben would rather be with Alice than anyone, and Bertie with Alex.

But Alice and Alex are already paired and effectively gone.

Bertie and Ben each prefer each other to all remaining players; they are 2nd order soulmates.

We refer to this process as *iterative matching of soulmates* [IMS] for the above players.

It these 4 were the only players the preference profile would be *IMS complete*.

This leaves Casey, Dannie, Eddie. No pair from these three are soulmates.

They would all prefer to have partners rather than be alone and have preferences as in our example above.

IMS is agnostic about the way in which a pair is formed from these three.

But in an SPNE two of them will be matched.

Proposition: The SPNE outcome satisfies IMS.

That is, all players who are soulmates are matched by IMS.

As shown in "Matching Soulmates", if IMS matches all players, the resulting outcome is the unique core coalition structure.

Corollary 1 *Suppose that all players are matched by IMS. Then every SPNE of an arbitrary ARG yields the unique core coalition structure.*

Moving on to rotating proposer games:

Rotating proposer games (RPGs): The order O is such that each i can make $|\mathbf{T}_i| + 1$ proposals before $i + 1$ can proposes.

Theorem 2. Every SPNE of an RPG is Pareto Optimal.

The proof begins with a series of observations.

Observation 1. For any strict subgame A, and for any two admissible proposals T; T', the resulting subgame if the same whether T or T' is rejected.★

Observation 2: Consider a proposer i and $k < |\mathbf{T}_i| + 1$, the k^{th} proposal node of player i.

Let π_{ik} be $\acute{i}$'s assigned coalition in the SPNE of the game that starts with her k^{th} proposal. Then either

$$\pi_{ik} \succ_i \pi_{i,k+1} \text{ or } \pi_{ik} = \pi_{i,k+1}.$$

.

Why? If $\pi_{i,k+1} \succ_i \pi_{ik}$, then in SPNE, when i proposes for the k^{th} time, she should make a proposal which will be rejected, contradicting Theorem 1. (In an SPNE all proposals are accepted). ★

Observation 3: If π_k is the SPNE outcome of the game beginning with player i's k^{th} proposal and

$$\pi_{ik} = \pi_{i,k+1},$$

then

$$\pi_k = \pi_{k+1}$$

.

The subgame that follows i proposing to π_{ik} and being accepted is the same whether it occurs following i's k^{th} or $k+1^{th}$ proposal.

Implying that if the next player j in O is not already in an coalition then the set of available players for j to propose to is the same in both cases.★

C**orollary.** There must be a $\overline{k}$ such that $\pi_{i\overline{k}} = \pi_{i,\overline{k}+1}$.

Lemma: Let π_1 be the SPNE outcome of a subgame A_1 with player i proposing for the first time, and let π_2 be the SPNE outcome of A_2, the subgame which results if i's first proposal is rejected. Then $\pi_1 = \pi_2$.

Proof:

1st: We show that if $\pi_{ik} = \pi_{i,k+1}$ then $\pi_{i,k-1} = \pi_{i,k+1}$.

The result then follows from Observation 3.

Assume $\pi_{ik} = \pi_{i,k+1}$.

From Observation 2, $\pi_{i,k-1} \succ \pi_{i,k}$ or $\pi_{i,k-1} = \pi_{i,k-}$. If $\pi_{ik} = \pi_{i,k-1}$ then by Observation 3, $\pi_{i,k-1} = \pi_{i,k-1}$ and the result follows as shown below.

Suppose instead, for contradiction, that $\pi_{i,k-1} \succ \pi_{i,k}$.

Since $\pi_{i,k-1}$ is accepted, we have that $\forall j \in \pi_{i,k-1}$, we have $\pi_{i,k-1} \succ \pi_{j,k}$.

Now since $\pi_{i,k-1} \succ \pi_{i,k}$ it must be that if i proposes to coalition $\pi_{i,k-1}$ on her k proposal, it is rejected.

(If $\pi_{i,k-1}$ were accepted in the SPNE, i would propose $\pi_{i,k-1}$ which would make i better off, contradicting the fact that π is an SPNE outcome of the game starting with i's k^{th} proposal.)

This implies that for some $j \in \pi_{i,k-1}$, it holds that $\pi_{i,k+1} \succ_j \pi_{i,k-1}$.

But since $\pi_{i,k} =_j \pi_{i,k+1}$ by our supposition, $\pi_{jk} \succ_j \pi_{i,k+1} =_j \pi_{jk}$ which gives us the contradiction.

Thus, $\pi_{i,k-1} = \pi_{ik}$.★

Proof of the Theorem 2. Let π be the SPNE outcome of an RPG.

For contradiction suppose π' is a Pareto-improvement over π.

Since $\pi \neq \pi'$, there are some players on different coalitions in π and π'.

Let i be the first such player.

All players in π_i and π'_i are still available when i first proposes.

Let A be the subgame starting with i's first proposal.

Since π' is a Pareto-improvement over π, from strict preferences it follows that $\pi'_i \succ_j \pi_j$ for all $j \in \pi'_i$.

If i's first proposal is to the coalition π_i', all members of the the coalition will accept. For otherwise they will be assigned, in the resulting SPNE, π_i'.

Therefore, i will propose to π_i' and the proposal will be accepted.

Implying π is not an SPNE outcome of the RPG, since the SPNE outcome is unique by Theorem 1.

We now have the desired contradiction, which concludes the proof.★

Conjecture: The number of proposals allowed by each player can be limited to 3 (or 2).

Why? Collectively, our results show that only one proposal, the last, is different than the prior proposals.

In summary,

- RPG equilibrium outcomes are
  - individually rational,
  - implement IMS (inherited from general ARGs),
  - Pareto optimal.
  - have no delay of forming coalitions and the SPNE outcome is unique.

- From the perspective of decentralized hedonic coalition formation with complete information, this is a strong set of properties.

The Rotating Proposer Mechanism

A *coalition formation mechanism* M maps every preference profile $\succ$ to a partition π, i.e. $\pi = M(\succ)$.

The Rotating Proposer Mechanism (RPM), implements the subgame perfect Nash equilibrium of the RPG in which all proposals are accepted.

For any profile, if all players report their preferences truthfully, equilibrium outcomes of the game have a number of good properties inherited by RPM.

Of particular note is that RPM is individually rational, Pareto optimal, and implements IMS.

However, from known results the RPM mechanism is not in general strategy-yproof.

For any profile, if all players report their preferences truthfully, equilibrium outcomes of the game RPG have a number of good properties which are thereby inherited by RPM.

In this case RPM is

(a) individually rational,

(b) Pareto optimal, and

(c) implements IMS.

Moreover, on the class of IMS complete preferences, RPM yields a unique core coalition structure.

These domains in which all players are incentivized to report their true preferences include well-known restrictions on preferences, such as top coalition and common ranking properties.

This, however, would seem to limit its practical consideration, as such restrictions can rarely be guaranteed or verified.

Moreover, we wish to make stronger efficiency claims than Pareto optimality, and also view fairness as an important criterion.

We are particularly interested in *utilitarian social welfare*, a much stronger criterion than Pareto efficiency. We will also consider several notions of fairness discussed below.

While we cannot make strong theoretical guarantees about these for broad realistic preference domains, we consider such properties empirically.

Empirical Methodology for the 3-Roomates Problem

It is computationally challenging to implement RPM.

The size of the backward induction search tree is $O(2^{\sum_{i \in N} |C_i|})$.

Even in roommate problem, computing SPE is $O(2^{n^2})$.

Summary of Empirical Results

Under RPM, incentives to misreport preferences are rare For the roommate problem, we find that typically 1% of players or fewer have an incentive to misrepresent their preferences, and fewer than 3% of all randomly generated profiles have *any* such players.

Approximate versions of RPM (which enable implementation of this mechanism at a larger scale) do not much degrade these results.

With coalitions of (at most) three, our experiment reflect only approximate RPM, and we find that no more than 5% of players have an incentive to misreport preferences.

RPM is highly efficient We compare utilitarian social welfare of RPM (using a cardinal transformation of ordinal preferences) and its approximations to serial dictatorship (which is also Pareto optimal).

We observe that in all experiments RPM yields much higher social welfare, with improvements typically ranging between 15 and 20%.

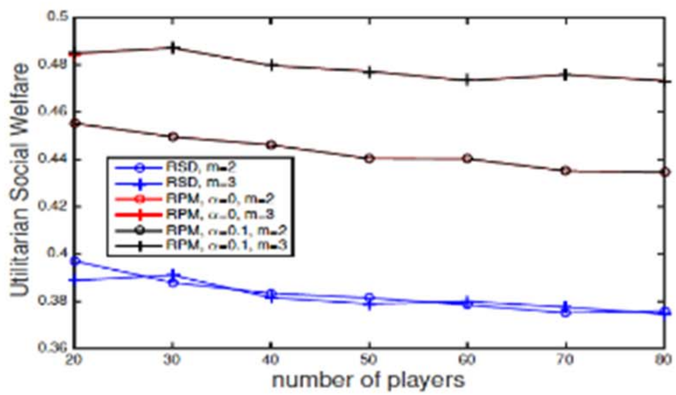
RPM yields equitable outcomes As is well known, serial dictatorship results in highly inequitable outcomes (in the ex post sense).

We observe that RPM yields outcomes far more equitable, with significant improvements in terms of the Gini coefficient, and a dramatically lower correlation between a random proposer order and utility (for example, correlation in some experiments drops from over 0.4 to well below 0.05).

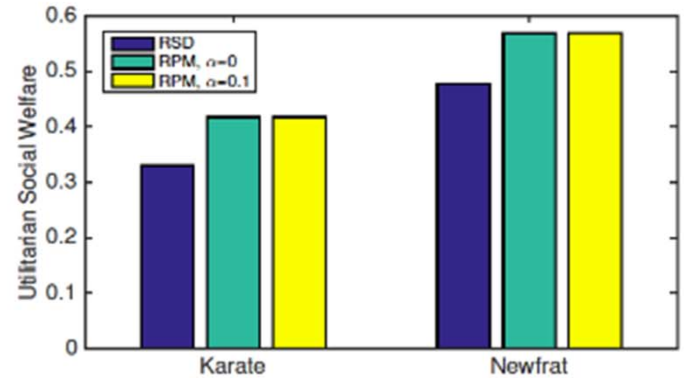
Remark: An important part of our empirical studies is that we can first run iterated matching of soulmates.

Just to give you an idea of more of our empirical results, the following charts illustrate how other mechanisms compare to ours in 4 different empirical studies.

Efficiency

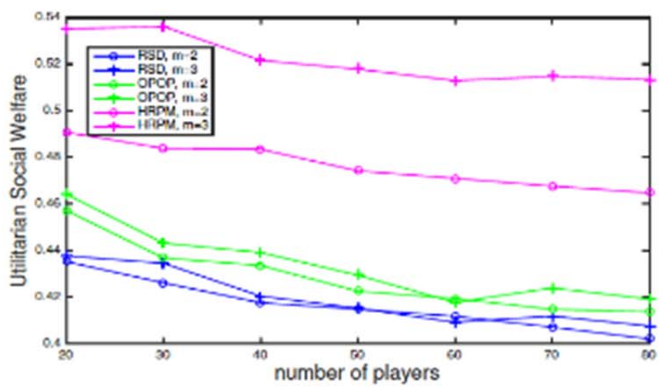


(a) scale-free networks

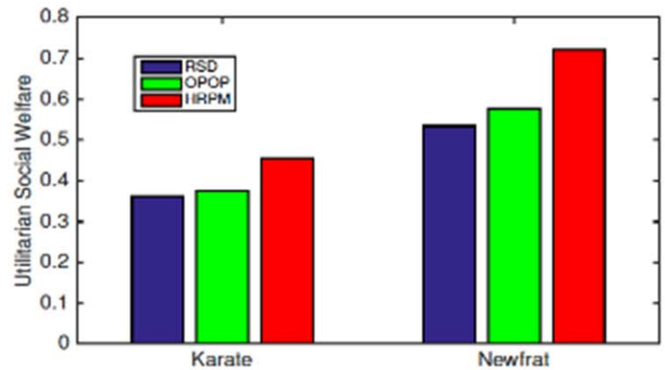


(b) Karate and Newfrat

Figure 6.3: Utilitarian social welfare for roommate problem



(a) scale-free networks



(b) scale-free networks

Figure 6.4: Utilitarian social welfare for trio-roommate problem

Conclusions

We consider sequential non-cooperative coalition formation games with a finite horizon.

Players iteratively propose coalitions, which are then sequentially accepted or rejected.

Our first key result is that there is an essentially unique no-delay equilibrium (all proposals are accepted in every equilibrium), and the equilibrium outcome is unique.

Our second major positive result is that in a subgame perfect Nash equilibrium coalitions involving soulmates, even in a stronger iterative sense, are always formed.

We also use it to provide a sufficient condition for the core outcome to be implemented in an equilibrium of our game.

Finally, we exhibit a restricted class of games, where the restriction is on the exogenously specified order of proposers, in which equilibrium outcomes are Pareto optimal.

Our most significant results demonstrate that the number of proposals a player can make affects the equilibrium outcome of the game.

These culminate in Theorem 2, which shows that with a sufficient number of proposals the equilibrium outcome is Pareto optimal, which is not the case if a player can only propose once.

This result is both novel and surprising. While the intuition—illustrated by an example—is that Pareto optimality results from the 1-proposal game serving as a credible threat, the proof is quite subtle and requires all the results obtained before this theorem.

While Theorem 2 is an interesting result for the class of coalition formation games considered, it also inspires a number of questions:

Most important, can similar results be obtained for other classes of games?

Are there other situations in which the ability of players to make multiple proposals can lead to Pareto improving outcomes?

We have in mind, in particular, political situations. The door is now open to the investigation of these, and other, related questions.

Figure 1:

Many thanks to the organizers and to all who attended this presentation.