

Bounding the Support Size in Extensive Form Games with Imperfect Information

Martin Schmid

Charles University In Prague
lifrordi@gmail.com

Matej Moravcik

Charles University In Prague
moravcikmatej42@gmail.com

Milan Hladik

Charles University In Prague
hladik@kam.mff.cuni.cz

Abstract

It is a well known fact that in extensive form games with perfect information, there is a Nash equilibrium with support of size one. This doesn't hold for games with imperfect information, where the size of minimal support can be larger. We present a dependency between the level of uncertainty and the minimum support size. For many games, there is a big disproportion between the game uncertainty and the number of actions available. In Bayesian extensive games with perfect information, the only uncertainty is about the type of players. In card games, the uncertainty comes from dealing the deck. In these games, we can significantly reduce the support size. Our result applies to general-sum extensive form games with any finite number of players.

Introduction

Arguably the most important solution concept in non-cooperative games is the notion of Nash equilibrium, where no player improves by deviating from this strategy profile. Support is defined as the set of actions played with non-zero probability and there are many crucial implications related to it.

Once the support is known, it is easy to compute the equilibrium in polynomial time even for general-sum games. Performance of some algorithms, namely the double-oracle algorithm for extensive form games, is tightly bound to the size of the support (Bosansky et al. 2013). Other work shows that minimizing the support in abstracted games can lead to better strategies in the original game (Ganzfried, Sandholm, and Waugh 2012). Finally, it is advantageous to prefer strategies having a small support. Such strategies are both easier to store and play.

Extensive form games model a wide class of games with a varying levels of uncertainty. In the case of perfect information, there is an optimal strategy using only one action in any information set. In contrast, in some extensive games with imperfect information, the player can be forced to use all the possible actions to play optimally.

In this paper, we focus on the relation between the level of uncertainty and the support size. We present an upper bound for the support size based on the uncertainty level.

Copyright © 2014, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Some games, such as Bayesian extensive games with observable actions or card games (such as no-limit Texas hold'em poker) have most of the information about the current state observable by all players, and therefore a low level of uncertainty. In these games, our bound guarantees the existence of Nash equilibrium having the support size considerably smaller than the number of all possible actions.

Instead of explicitly defining a level of uncertainty, we use the concept of the *public tree*. This concept provides a nice interpretation of uncertainty and *public actions*. Using the public tree, we present a new technique called the *equilibrium preserving transformation*, which transforms some equilibrium strategy profile into another. We provide an upper bound on the number of public actions used in the transformed Nash equilibrium.

Our approach also applies to games with non-observable actions, where it simply limits the number of public actions.

Applying our result to specific games, we present a new bound for the support size in these games.

For example, in no-limit Texas hold'em poker, there can be any finite number of actions available in some information sets. Our result implies the existence of an optimal strategy for which the number of actions used in every information set depends only on the number of players and the number of card combinations players can be dealt.

In Bayesian extensive games with observable actions, the bound equals to the number of different player types the chance can reveal.

Moreover, our proof is constructive. Given an extensive form game and an optimal strategy, the equilibrium preserving transformation finds another optimal strategy satisfying our bound in polynomial time.

Background

Extensive form games (Osborne and Rubinstein 1994, p. 200). An extensive form game consists of

- A finite set N (the set of **players**).
- A finite set H of sequences. Each member of H is a **history**, each component of history is an **action**. The empty sequence is in H , and every prefix of a history is also history ($((h, a) \in H \implies (h \in H))$). $h \sqsubseteq h'$ denotes that h is a prefix of h' . $Z \subseteq H$ are the terminal histories (they are not a prefix of any other history).

- The set of actions available after every non-terminal history $A(h) = \{a : (h, a) \in H\}$.
- A function p that assigns to each non-terminal history an **acting player** (member of $N \cup c$, where c stands for chance).
- A function f_c that associates with every history for which $p(h) = c$ a probability measure on $A(h)$. Each such probability measure is independent of every other such measure.
- For each player $i \in N$, a partition \mathcal{I}_i of $h \in H : p(h) = i$. \mathcal{I}_i is the **information partition** of player i . A set $I_i \in \mathcal{I}_i$ is an **information set** of player i .
- For each player $i \in N$ an **utility function** $u_i : Z \rightarrow \mathbb{R}$.

We assume that the game satisfies the **perfect recall**, which guarantees that the players never forget any information revealed to them nor the order in which the information was revealed. Games that do not satisfy this property are not guaranteed to have a behavioral Nash equilibrium.

Public Game Tree

The essential concept in our approach is that of a public game tree. Informally, a public game tree is a game view for an observer that knows no private information.

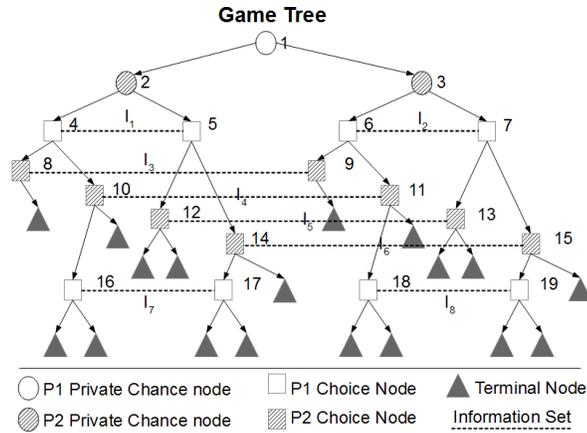


Figure 1: Extensive form game tree for one-card poker.

Public game tree, as introduced in (Johanson et al. 2011) is a partition of the histories. P is a **public partition** and $\rho \in P$ is a **public state** if

- No two histories in the same information set are in different public states.
- Two histories in different public states have no descendants in the same public state.
- No public state contains both terminal and non-terminal histories (public states are either terminal or non-terminal).

The public tree offers a nice interpretation of imperfect information. For games with perfect information, the public tree is the same as the game tree. As the uncertainty grows, more and more information sets collapse into a single public state.

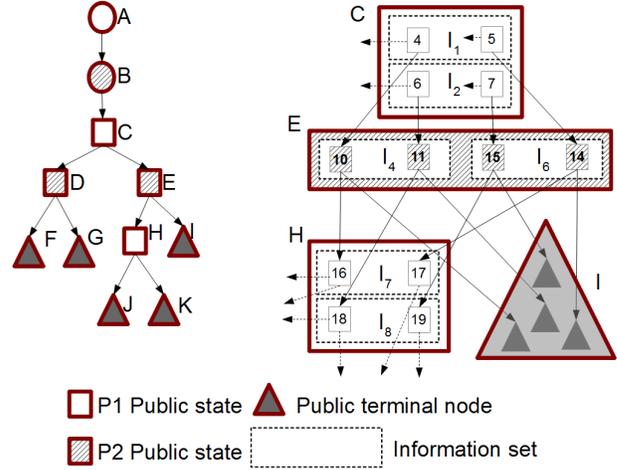


Figure 2: The public tree of a game in Figure (1). The structure of public states C, E, H, I is:

$$\begin{aligned} \text{prev}(E, 1) &= \{I_1, I_2\}, \text{last}(E) = \{I_4, I_6\} \\ C(E) &= \{H, I\}, A(I_1 \rightsquigarrow E) = [4] \rightarrow [10] \\ A_p(C \rightsquigarrow E) &= \{(I_1, [4] \rightarrow [10]), (I_1, [5] \rightarrow [14]), \\ & (I_2, [6] \rightarrow [11]), (I_2, [7] \rightarrow [15])\} \end{aligned}$$

For the public game tree, we also define:

- The set of acting players in $\rho \in P$ as $p(\rho)$.
- If the same player acts in all histories $h \in \rho$, then we define the acting player in ρ as: $p(\rho) = p(h)$ for some $h \in \rho$.
- $C(\rho)$ to be set of child public states of ρ .
- For any public state ρ and any player i , we define $\text{prev}(\rho, i)$ to be the set of player's last information sets he could play in before reaching ρ . $I \in \text{prev}(\rho, i)$ if:

$$\begin{aligned} p(I) &= i, \text{ there are histories } h \in \rho, h' \notin \rho, h' \in I: \\ & h' \sqsubseteq h \text{ and there is no history } h'' \notin \rho : h'' \sqsubseteq h, \\ & h' \sqsubseteq h'', p(h'') = i \end{aligned}$$

- For any public state ρ , we define $\text{last}(\rho)$ to be the last information sets the player $p(\rho)$ plays before leaving ρ . $I \in \text{last}(\rho)$ if $I \in \rho$, there is some history $(h, a) \notin \rho$ and $h \in \rho$.
- We call a non-terminal public state ρ **simple** if an observer knows which player acts and there are no information sets that contain actions going to a different public state as well as the actions that aren't. Formally, $p(h) = p(h') \neq c$ for all $h, h' \in \rho$ and no history has prefix in two different information states from $\text{last}(\rho)$.
- We define $A(I \rightsquigarrow \rho)$ to be actions that a player can take in $I \in \mathcal{I}$ in order to get to the public state ρ :

$$A(I \rightsquigarrow \rho) = \{a \in A(I) | h \in I, (h, a) \sqsubseteq h', h' \in \rho\}$$

- Intuitively, actions are the edges connecting any two nodes (histories) in Figure (1). We want **public actions** to be the edges connecting any two public states as seen on Figure (2). We define the public action going from public state ρ to ρ' as a set of pairs (*information set, action*):

$$A_p(\rho \rightsquigarrow \rho') = \{(I, a) | I \in \text{last}(\rho), a \in A(I \rightsquigarrow \rho')\}$$

See the box under the Figure (2) for examples of these definitions.

Strategies and Equilibrium

Strategies and Utility

A **strategy** for player i , σ_i , is a function that maps $I \in \mathcal{I}_i$ to a probability distribution over $A(I)$ and $\pi^\sigma(I, a)$ is the probability of action a . Σ_i denotes the set of all strategies of player i . A **strategy profile** is a vector of strategies of all players, $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_{|N|})$. Σ denotes the set of all strategy profiles.

We denote $\pi^\sigma(h)$ as the probability of history h occurring given the strategy profile σ . Let $\pi_i^\sigma(h)$ be the contribution of player i to that probability. We can then decompose $\pi^\sigma(h)$ as

$$\pi^\sigma(h) = \prod_{i \in N \cup c} \pi_i^\sigma(h)$$

Let $\pi_{-i}(h)$ be the product of all players contribution (including chance), except that of player i . Define $\sigma|_{I \rightarrow a}$ to be the same strategy profile as σ , except that a player always plays the action a in the information set I . Define $u_i(\sigma)$ to be the expected utility for player i , given the strategic profile σ .

The **support** of a strategy profile σ , $support^\sigma(I)$, is the set of actions that the player plays with non-zero probability in I .

We say that the public action $A_p(\rho \rightsquigarrow \rho')$ is **supported** given the strategy profile σ , if for any $(I, a) \in A_p(\rho \rightsquigarrow \rho')$, $a \in support^\sigma(I)$. Let $support_p^\sigma(\rho)$ be the set of all supported public actions in ρ :

$$support_p^\sigma(\rho) = \{A_p(\rho \rightsquigarrow \rho') | A_p(\rho \rightsquigarrow \rho') \text{ is supported in } \sigma\}$$

Nash Equilibrium and Regret

A **Nash equilibrium** is a strategy profile σ such that for any player $i \in N$,

$$u_i(\sigma) \geq \max_{\sigma_i^* \in \Sigma_i} u_i(\sigma_i^*, \sigma_{-i})$$

Overall regret R_i^σ is the difference between the player's utility given the strategy profile σ and the single strategy that would maximize his value:

$$R_i^\sigma = \max_{\sigma_i^* \in \Sigma_i} u_i(\sigma_i^*, \sigma_{-i}) - u_i(\sigma)$$

Clearly, R_i is always non-negative and there is a simple relation between the overall regret and Nash equilibrium:

$$R_i^\sigma = 0 \forall i \in N \iff \sigma \text{ is Nash equilibrium}$$

Counterfactual Values, Regret and Equilibrium

To show that some strategy profile σ is an equilibrium, we could show that the regret $R_i^\sigma = 0$ for all players.

There is a way to bound this *full* regret R_i^σ using *partial* regrets in all information sets. These partial regrets are called counterfactual regrets.

The **Counterfactual utility** $u_i(\sigma, I)$ is the expected utility given that information set I is reached and all players play using strategy σ , except that player i plays to reach I (Zinkevich et al. 2007):

$$u_i(\sigma, I) = \frac{\sum_{h \in I, h' \in Z} \pi_{-i}^\sigma(h) \pi^\sigma(h'|h) u_i(h')}{\pi_{-i}^\sigma(I)}$$

The **Counterfactual regret** (Zinkevich et al. 2007) is then defined as

$$R_i^\sigma(I) = \max_{a \in A(I)} \pi_{-i}^\sigma(I) (u_i(\sigma|_{I \rightarrow a}, I) - u_i(\sigma, I))$$

Theorem 1. (Zinkevich et al. 2007)

$$R_i^\sigma \leq \sum_{I \in \mathcal{I}_i} R_i^\sigma(I)$$

Note that in contrast to (Zinkevich et al. 2007), we are not interested in the relation between average regret and ϵ -equilibrium (which holds only for two players, zero-sum games). We are interested only in bounding the regret of strategy profile σ using the Theorem (1). Directly from that theorem, we get the following corollary

Corollary 1. *If $R_i^\sigma(I) = 0$ for all $I \in \mathcal{I}_i, i \in N$, the strategy profile σ forms a Nash equilibrium.*

The converse implication is not true in general. There can be an equilibrium with $R_i^\sigma(I) > 0$ for some $I \in \mathcal{I}_i$. But there is always some Nash equilibrium for which $R_i^\sigma(I) = 0$ for all $I \in \mathcal{I}_i$. (It is easy too see that if $R_i^\sigma(I) > 0$ for some I , the player i plays not to reach I , $\pi_i^\sigma(I) = 0$. One can make the counterfactual regret zero in these sets using backward induction.)

Finally, there is a simple way to show that the strategy profile σ' is a Nash equilibrium by comparing it with another Nash equilibrium.

Lemma 1. *Given a Nash equilibrium σ , for which $\sum_{I \in \mathcal{I}_i} R_i^\sigma(I) = 0$, if we find a strategy profile σ' such that for all $i \in N, I \in \mathcal{I}_i$ and $A \in A(I)$*

$$u_i(\sigma|_{I \rightarrow a}, I) = u_i(\sigma'|_{I \rightarrow a}, I)$$

and in every information set, the strategy σ'_i assigns a non zero probability only to actions used with non-zero probability in σ , the strategy profile σ' forms a Nash equilibrium for which $\sum_{I \in \mathcal{I}_i} R_i^{\sigma'}(I) = 0$.

The proof follows directly from the definition of counterfactual regret.

Main Theorem

Our main result shows the existence of optimal strategy with a limited number of supported public actions in simple public states.

Theorem 2. *In any finite extensive form game, there is an equilibrium strategy profile σ such that for every simple public state ρ , the number of supported public actions, $|\text{support}_p^\sigma(\rho)|$, is bounded by*

$$|\text{last}(\rho)| + \sum_{j \in N \setminus \{p(\rho)\}} \sum_{I \in \text{prev}(\rho, j)} |A(I \rightsquigarrow \rho)|$$

This bound has a nice interpretation for some specific games. For example in games with publicly observable actions, the bound depends only on the uncertainty presented by the chance.

We limit our technique only to simple public states, but it is possible to generalize the result. If there are some information sets containing actions going to a different public state as well as the actions that don't, it's easy to come up with an equivalent game where all actions are public. We are not aware of any well-known extensive form game having public states that are not simple.

Overview of our Approach

The core of our approach is a new technique we call **equilibrium preserving transformation**. We start with a Nash equilibrium for which $R_i^\sigma(I) = 0$ for all $i \in N$ and $I \in \mathcal{I}_i$, which is guaranteed to exist. An equilibrium preserving transformation carefully shifts probabilities locally, using the public tree. Given a public state ρ , it shifts some probabilities in information sets in ρ . The point is to keep the strategy optimal for all players, while minimizing the number of supported public actions.

Applying this transformation to a single public state ρ , we get a new equilibrium where the number of supported public actions satisfies our bound. Thus, we bound the public actions used in that information set, but we don't touch the strategies in any other public state.

Applying this transformation again to the new equilibrium, but in a different public state, we bound the number of supported public actions in that public state. Since we don't touch actions in any other public state, we do not violate the bound from the previous step.

Repeating this for all public states, we finally get a Nash equilibrium where the bound holds for all simple public states.

Optimality of the New Strategy

To show that the new strategy is an equilibrium, we leverage the concept of counterfactual values. These values are defined at the level of information sets and we can show that the strategy is optimal by showing that these values remain unchanged thanks to Lemma (1). Since we change the strategy only in the information sets that are highly structured (they are in the same public state), it's relatively easy to compute the changed counterfactual values in all information sets.

Equilibrium Preserving Transformations

Given a Nash equilibrium σ where $R_i^\sigma(I) = 0$ for all $I \in \mathcal{I}$, the core idea of our approach is to transform this strategy

profile to another strategy profile σ' . We refer to this transformation as equilibrium preserving transformation or EPT and we denote the transformed strategy as $\sigma' = EPT(\sigma)$.

EPT shifts probabilities for a player locally, using the public tree. Given some public state ρ , we carefully change probabilities of outgoing actions. Since there's only one player acting in ρ , we change strategy only for this player $p(\rho) = i$.

We change the strategy only in $\text{last}(\rho)$, which are the last information sets in which the player acts just before reaching some $\rho' \in C(\rho)$.

We will continuously add some restrictions to our transformation and show what these restrictions imply for the new strategy profile σ' . Finally, we will see that if we transform the strategy such that all these restrictions hold, σ' is a Nash equilibrium.

Restriction 1 The transformed strategy profile σ' differs from σ only in information sets $I \in \text{last}(\rho)$

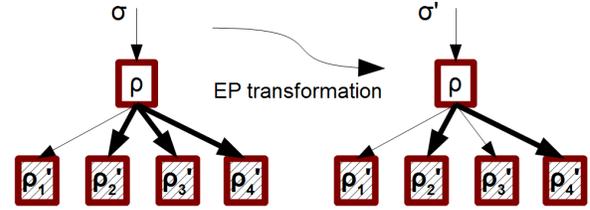


Figure 3: In the new strategy σ' , we bound the number of supported public actions (bold arrows).

The trick is to shift the probabilities in these information sets to use as few actions as possible, while keeping the strategy profile equilibrium. To do that, we keep the counterfactual values unchanged for all information sets and all actions.

To insure this, we impose two restrictions. The first one fixes counterfactual values for all information sets *after* ρ . The second one (together with the first one) fixes counterfactual values for all other information sets.

Information Sets after ρ

To fix the counterfactual values for the information sets after ρ , we do not shift the strategies arbitrarily. We only multiply some action probabilities with carefully chosen constants.

The last information sets the player $p(\rho)$ acts in before leaving ρ are $\text{last}(\rho)$. We consider the probability of all actions $a \in A(I \rightsquigarrow \rho')$ for any $I \in \text{last}(\rho)$. Our transformation is only allowed to multiply these action probabilities using some constant, which we call $\kappa(\rho \rightsquigarrow \rho')$.

Restriction 2 The probabilities of all outgoing actions from ρ to some child public state ρ' are all multiplied with some constant $\kappa(\rho \rightsquigarrow \rho')$. For all $I \in \text{last}(\rho)$ and $a \in A(I \rightsquigarrow \rho')$

$$\pi^{\sigma'}(I, a) = \kappa(\rho \rightsquigarrow \rho') \pi^\sigma(I, a) \quad (1)$$

We show later how to ensure that once we multiply the probabilities with corresponding κ , we get a valid strategy in every information set.

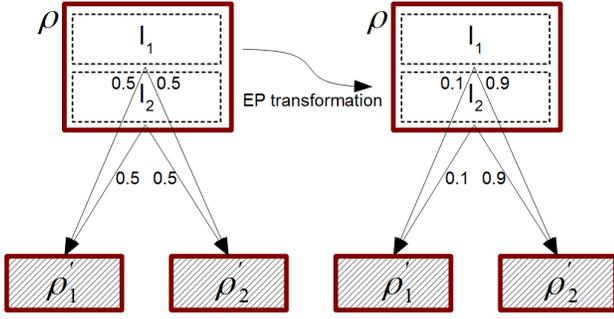


Figure 4: Transformation multiplies probabilities of all actions that go from ρ to some ρ' with corresponding κ . In this case $\kappa(\rho \rightsquigarrow \rho'_1) = \frac{1}{5}$ and $\kappa(\rho \rightsquigarrow \rho'_2) = \frac{9}{5}$.

The reason why the transformation is not allowed to change the probabilities arbitrarily, but can only multiply action probabilities of actions going from ρ to ρ' with some corresponding $\kappa(\rho \rightsquigarrow \rho')$, is to keep the counterfactual values in all information sets after ρ unchanged.

Lemma 2. *If $\sigma' = EPT(\sigma, \rho)$, counterfactual values in all information sets after ρ remain unchanged.*

$u_j(\sigma'|_{I \rightarrow a}, I) = u_j(\sigma|_{I \rightarrow a}, I)$ for all $j \in N$ and for all I after ρ .

The proof (in Appendix A) follows directly from the definition of $u_j(\sigma'|_{I \rightarrow a}, I)$.

Multiplying the strategies with corresponding κ is the only transformation we do. Clearly, if $\kappa(\rho \rightsquigarrow \rho') = 0$ for some ρ' , $A_p(\rho \rightsquigarrow \rho')$ is not supported in the new strategy profile. In other words, we are interested in finding as many zero variables κ as possible.

Other Information Sets

The previous constraints keep the counterfactual values unchanged for all players and all information sets after ρ .

To ensure that the counterfactual values don't change in other information sets, we shift strategies such that the counterfactual values are unchanged in every $I \in \text{prev}(\rho, j)$ for all players $j \neq p(\rho)$.

Restriction 3 For all players $j \neq p(\rho)$, for all $I \in \text{prev}(\rho, j)$ and for all $a \in A(I \rightsquigarrow \rho)$

$$u_j(\sigma'|_{I \rightarrow a}, I) = u_j(\sigma|_{I \rightarrow a}, I) \quad (2)$$

The point is that if we keep counterfactual values unchanged only in these information sets, counterfactual values in all other information sets remain the same.

Lemma 3. *If $\sigma' = EPT(\sigma, \rho)$, counterfactual values in all information sets remain unchanged.*

For the proof see Appendix A.

System of Linear Equations to Find κ

Now we will show how to find $\kappa(\rho \rightsquigarrow \rho')$ such that all restrictions are satisfied. We are interested only in ρ' such

that $A_p(\rho \rightsquigarrow \rho')$ is supported, all other actions have zero probability anyway. We find all κ using a systems of linear equations (linearity is the crucial part), with variables $\kappa \geq 0$.

First set of equations makes sure that σ' is a valid strategy. Adding another set of equations ensures that the counterfactual values remain unchanged.

Finally, using the simple property of linear equations, there must be a basic solution having limited number of non-zero variables κ .

First System of Equations

First, we write a simple equation for every $I \in \text{last}(\rho)$ to make sure that we get a valid strategy after multiplying with corresponding κ

$$\sum_{\rho' \in C(\rho)} \sum_{a \in A(I \rightsquigarrow \rho')} \pi^\sigma(I, a) \kappa(\rho \rightsquigarrow \rho') = 1 \quad (3)$$

Since we write down this equation for every $I \in \text{last}(\rho)$, there are $|\text{last}(\rho)|$ of equations in total. Note that these equations are indeed linear in the variable κ .

Second System of Equations

The second system of linear equations makes sure that the restriction (3) is satisfied.

First, we compute the counterfactual values for the strategy profile σ' using the variables κ and the strategy profile σ .

Lemma 4. *There are some constants $c_0(I, a) \dots c_{|C(\rho)|}(I, a)$ such that the counterfactual utility for all players $j \neq p(\rho)$, for all $I \in \text{prev}(\rho, j)$ and for all $a \in A(I \rightsquigarrow \rho)$*

$$u_j(\sigma'|_{I \rightarrow a}, I) = c_0(I, a) + \sum_{i=\{1 \dots |C(\rho)|\}} \kappa(\rho \rightsquigarrow \rho'_i) c_i(I, a)$$

For the proof see Appendix A.

In the proof of the above lemma, we leverage the fact that the values are unchanged in all information sets $\rho' \in \rho$. The history either passes through some ρ' and its probability gets multiplied with corresponding κ , or it doesn't and the probability remains unchanged.

Using this result, we can simply add linear equations for all players $j \neq p(\rho)$, for all $I \in \text{prev}(\rho, j)$ and for all $a \in A(I \rightsquigarrow \rho)$

$$\sum_{i=\{1 \dots |C(\rho)|\}} \kappa(\rho \rightsquigarrow \rho'_i) c_i(I, a) = u_j(\sigma|_{I \rightarrow a}, I) - c_0 \quad (4)$$

Again, these equations are linear in the variable κ .

Final System of Equations

Putting together all equations from (3) and (4), we are interested in $\kappa \geq 0$ such that

$$\forall I \in \text{last}(\rho)$$

$$\sum_{\rho' \in C(\rho)} \sum_{a \in A(I \rightarrow \rho')} \pi^\sigma(I, a) \kappa(\rho \rightarrow \rho') = 1$$

$$\forall j \in N \setminus \{p(\rho)\}, I \in \text{prev}(\rho, j), a \in A(I)$$

$$\sum_{i=\{1 \dots |C(\rho)|\}} \kappa(\rho \rightsquigarrow \rho'_i) c_i(I, a) = u_j(\sigma|_{I \rightarrow a}, I) - c_0 \quad (5)$$

Combining previous results, we get a straightforward corollary

Corollary 2. *Any solution to (5) defines a valid equilibrium preserving transformation.*

This polyhedron is clearly bounded and non-empty ($\kappa = 1$ is a solution to (5)). Finally, we use the well known property of basic solutions. There must be some basic solution where the number of non-zero variables κ is no larger than the number of equations (Bertsimas and Tsitsiklis 1997). Since number of equations is

$$|\text{last}(\rho)| + \sum_{j \in N \setminus \{p(\rho)\}} \sum_{I \in \text{prev}(\rho, j)} |A(I \rightsquigarrow \rho)| \quad (6)$$

our main theorem is proven. Moreover, we can find this solution efficiently in polynomial time using linear programming (Ye 1991).

Example Games

In this section, we mention few existing games and show how our bound applies to these. As far as we know, these are the first bounds on the support size presented for these games.

Games where the players see the actions of all other players are called games with publicly observable actions. The only uncertainty comes from the actions of chance. In these games, all public states are simple and all information sets in any public state ρ form $\text{last}(\rho)$.

Because all actions are public, $|A(I \rightsquigarrow \rho)| = 1$ for all ρ and for all $I \in \text{prev}(\rho, j)$. Consequently, the second term of (6) becomes $\sum_{j \in N \setminus \{p(\rho)\}} |\text{prev}(\rho, j)|$.

Finally, the bound for supported public actions implies an upper bound on the size of $\text{support}^\sigma(I)$ for every $I \in \rho$.

$$|\text{support}^\sigma(I)| \leq |\text{support}_p^\sigma(\rho)|$$

Bayesian Extensive Games with Observable Actions

Bayesian extensive games with observable actions (Osborne and Rubinstein 1994, p. 231) are games with publicly observable actions, where the only uncertainty comes from the initial move of chance. Chance selects a player type $\theta \in \Theta_i$ for each player i . Because chance acts at the very beginning of the game, the number of information sets grouped in every public state ρ , equals $|\Theta_{p(\rho)}| = |\text{last}(\rho)|$. Similarly, $|\text{prev}(\rho, j)| = |\Theta_j|$.

Corollary 3. *For any Bayesian extensive games with observable actions, there's a Nash equilibrium where the size of $\text{support}^\sigma(I)$ for any $I \in \mathcal{I}$ is bounded by*

$$\sum_{i \in N} |\Theta_i| \quad (7)$$

No-limit Texas Hold'em Poker

In Texas hold'em poker, players are dealt two private cards out of a deck of 52 cards. Four betting rounds follow and dealer deals some more publicly visible cards between these betting rounds.

In no-limit version, players can bet any amount of money up to their stack in every betting round. For example in the 2014 AAI Computer Poker Competition, there are up to 20000 actions available in information sets (ACPC 2014).

All betting is publicly visible, and the only uncertainty is about the private cards the players were dealt. $|\text{prev}(\rho, j)| = |\text{last}(\rho)| = \binom{52}{2}$ for any player j and any public state ρ .

Corollary 4. *In Texas hold'em poker, there's a Nash equilibrium where the size of $\text{support}^\sigma(I)$ for any $I \in \mathcal{I}$ is bounded by*

$$\binom{52}{2} |N| \quad (8)$$

Using some isomorphisms, we can further decrease the bound in some situations. In Texas hold'em poker, there are 169 non-isomorphic (Waugh 2013) pairs in the first round (called *preflop*).

Corollary 5. *In Texas hold'em poker, there's a Nash equilibrium where the size of $\text{support}^\sigma(I)$ for all information set in the first round (*preflop*) is bounded by*

$$169 |N| \quad (9)$$

Conclusion

We present a new technique called equilibrium preserving transformation, that implies the existence of Nash equilibrium having a bounded support. Our bound shows a relation between the level of uncertainty and the support size.

For Bayesian extensive games with observable actions and card games, our bound implies relatively small support. Finally, given any Nash equilibrium, EPT finds another equilibrium having the bounded support in polynomial time.

Future Work

We intent to explore the possibility of using EPT for finding good abstractions of action space in games such as poker. The idea is to first create a toy game with small level of uncertainty. Once we find Nash equilibrium in this toy game, we can use EPT to get an equilibrium using only a small number of actions. Hopefully, this approach will help us to find actions that also work nicely in the original game.

Acknowledgment

The work was supported by the Czech Science Foundation Grant P402/13-10660S and by the Charles University Grant SVV-2014-260103.

References

- ACPC. 2014. Acpc 2014 rules, computerpokercompetition.org.
- Bertsimas, D., and Tsitsiklis, J. N. 1997. Introduction to linear optimization.

Bosansky, B.; Kiekintveld, C.; Lisy, V.; Cermak, J.; and Pechoucek, M. 2013. Double-oracle algorithm for computing an exact nash equilibrium in zero-sum extensive-form games. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, 335–342. International Foundation for Autonomous Agents and Multiagent Systems.

Ganzfried, S.; Sandholm, T.; and Waugh, K. 2012. Strategy purification and thresholding: Effective non-equilibrium approaches for playing large games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, 871–878. International Foundation for Autonomous Agents and Multiagent Systems.

Johanson, M.; Waugh, K.; Bowling, M.; and Zinkevich, M. 2011. Accelerating best response calculation in large extensive games. In *Proceedings of the Twenty-Second international joint conference on Artificial Intelligence-Volume One*, 258–265. AAAI Press.

Osborne, M. J., and Rubinstein, A. 1994. *A course in game theory*. MIT press.

Waugh, K. 2013. A fast and optimal hand isomorphism algorithm. In *Workshops at the Twenty-Seventh AAAI Conference on Artificial Intelligence*.

Ye, Y. 1991. An $O(n^3)$ potential reduction algorithm for linear programming. *Mathematical programming* 50(1-3):239–258.

Zinkevich, M.; Johanson, M.; Bowling, M.; and Piccione, C. 2007. Regret minimization in games with incomplete information. In *Advances in neural information processing systems*, 1729–1736.