Computing all the mixed-strategy equilibria in the repeated prisoner’s dilemma

Kimmo Berg
Department of Mathematics and Systems Analysis
Aalto University, Finland

July 18, 2018
Outline of the presentation

- Comparison of pure, correlated pure and mixed equilibria
  - Result 1: Mixed payoff set is dramatically different
  - Result 2: Max payoff higher, more Pareto points

- Repeated games are solved using set-valued games

- How to solve set-valued games?
  - Decompose into classes and solve each class separately
  - Result 3: Sufficient to compute extreme points
    - X-Y convexity
    - Monotonicity property
The model

- Infinitely repeated two-player game
- Stage game with finitely many actions
- Discounting with possibly unequal discount factors
- Behavior strategies: randomization and history-dependent
- Players observe realized pure actions, not randomizations
The model (2)

- Finite set of players \( N = \{1, \ldots, n\} \)
- Finite set of pure actions \( A_i, \ i \in N, \ A = \times_{i \in N} A_i \)
- Mixed action \( q_i(a_i) \geq 0 \), profile \( q = (q_1, \ldots, q_N) \)
- Probability of pure-action profile \( a \in A \): \( \pi_q(a) = \prod_{j \in N} q_j(a_j) \)
- Stage-game payoff \( u_i(q) = \sum_{a \in A} u_i(a) \pi_q(a) \)
- Histories \( H^k = A^k \) for stage \( k \geq 0 \), \( H^0 = \emptyset \)
- Behavior strategy \( \sigma_i : H \mapsto Q_i \)
- Discounted payoff \( U_i(\sigma) = \mathbb{E} \left[ (1 - \delta_i) \sum_{k=0}^{\infty} \delta_i^k u_i^k(\sigma) \right] \)
Prisoner’s Dilemma

<table>
<thead>
<tr>
<th></th>
<th>(a)</th>
<th>(b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5,3.5</td>
<td></td>
<td>0,4</td>
</tr>
<tr>
<td>4,0</td>
<td>1,1</td>
<td></td>
</tr>
</tbody>
</table>

What are equilibria in pure, correlated and mixed strategies?

- Common discount factor $\delta$
- Pure-action profiles are called $a$, $b$, $c$ and $d$
- Correlated pure: APS (1990), Judd et al. (2003), AS (2014)
Impatient players: $0 \leq \delta < 1/6 \approx 0.167$

![Graphs showing pure, correlated pure, and mixed strategies for impatient players.]

Impatient players can only play the Nash equilibrium.
Low discounting: $1/6 \leq \delta < 2/7 \approx 0.286$

Result 1: Set of mixed equilibrium can be dramatically different
Mixed-equilibrium payoffs for $\delta = 1/4$

- Close to convergence: area changes less than $3 \cdot 10^{-8}$
- At least approximate equilibria
Medium discounting: \( 2/7 \leq \delta < 8/13 \approx 0.615 \)

Result 2: Maximum payoff higher and more Pareto points
Mixed-equilibrium payoffs for $\delta = 4/10$

- Area 6.83 is close to the FIR area $7\frac{1}{7} \approx 7.14$.
- No convergence problems here.
High discounting: $\delta \geq \frac{8}{13} \approx 0.615$

All the feasible and individual rational (FIR) payoffs are obtained (Folk theorem). Note that the equilibrium payoffs need not be inside FIR for unequal discount factors. See Berg and Karki (2018): Critical discount factor values in discounted supergames.
Set-valued game

- Payoffs are chosen from sets: $K_x$ for action profile $x$
- Tuple $G' = (N, A, K)$, $K = \times_{x \in A} K_x$
- Nash equilibria of a set-valued game is

$$M(G') = \bigcup \{ M(u(x)) : u(x) \in K_x, x \in A \}.$$  

- Useful in other game models (e.g., uncertain payoffs)
Two set-valued games and their equilibria

(a) (b)
Repeated games are solved using set-valued games

- The payoff set $V$ is the largest fixed-point of mapping $B$:

  $$W = B(W) \doteq \bigcup_{z(x) \in W} M((I - T)u(x) + Tz(x)).$$

- $M(y)$ is the Nash equilibria in stage game with payoffs $y$
- $T$ is diagonal matrix of discount factors
- $u(x)$ is the stage-game payoff of action profile $x$
- $z(x)$ is the continuation payoff after action profile $x$
Iteration of $V_{i+1} = B(V_i)$

Initial guess: $V_0$ is chosen as the square from (1, 1) to (4, 4)
Each iteration corresponds to solving a set-valued game
Split set-valued game into classes of games

<table>
<thead>
<tr>
<th></th>
<th>3.51,3.51</th>
<th>0.88,3.39</th>
<th>3.39,3.51</th>
<th>0.25,3.25</th>
<th>3.25,0.41</th>
<th>1,1</th>
<th>3.39,0.88</th>
<th>1,1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two mixed-strategy classes: C14 (black dots) and C12 (dashed)</td>
<td>3.51,3.51</td>
<td>0.88,3.39</td>
<td>3.39,3.51</td>
<td>0.25,3.25</td>
<td>3.25,0.41</td>
<td>1,1</td>
<td>3.39,0.88</td>
<td>1,1</td>
</tr>
</tbody>
</table>
C14 is the hardest class with truly mixed strategies

- Enough to compute extreme points and take X-Y convex hull
- Relies on monotonicity result
- North-east frontier is shown on the right and the 16 points
Truly mixed-strategy payoff

\[ v_i = \frac{a_id_i - b_ic_i}{a_i + d_i - b_i - c_i} \]

- Nonlinear but continuous and monotone in class
- Payoffs \( a_i \) to \( d_i \) do not directly affect \( j \)'s payoff \( v_j \)
- Indirect effect due to sets \( K_x \) since \((a_i, a_j) \in K_a\) are coupled
Monotonicity results

**Definition**
A set $S$ is X-Y convex if it is connected and contains all the vertical and horizontal lines.

**Proposition**
*In 2 × 2 set-valued games without ties, the truly mixed-strategy equilibrium payoffs are monotone in the players’ payoffs if the signs of $b_i - d_i$, $c_i - d_i$, $c_i - a_i$, and $b_i - a_i$, $i = 1, 2$, remain the same.*

**Proposition**
*In a 2 × 2 set-valued game, if the sets $K_a - K_d$ are X-Y convex and the payoffs are monotone, then the extreme points of the truly mixed-strategy equilibrium payoffs are produced by the extreme points of the X-Y convex sets.*
Class C12 (or C4) is easier to solve

- Only need to find the southern border
- Traces maxima of the sets $K_a$ and $K_c$
Classes C1 and C4 (darker shade)
No need to compute truly mixed strategies (C12 and C14)
Those payoffs will be inside the C1 rectangles
Repeated games are solved using set-valued games
Set-valued games are split into classes
Each class has specialized algorithm
Typically, only few classes need to be solved
It is enough to compute certain extreme points
Sets have to be split into X-Y convex parts for truly mixed strategies
That’s all folks...

Thank you! Any questions?