Best Experienced Payoff Dynamics and Cooperation in the Centipede Game and the Traveler's Dilemma

William H. Sandholm
University of Wisconsin

Segismundo S. Izquierdo
Universidad de Valladolid

Luis R. Izquierdo
Universidad de Burgos

Introduction

Backward induction and rationalizability are pillars of game theory.

They are defined by iterative procedures rather than equilibrium logic,
but still provide tight predictions in some interesting games,
and are the bases for elegant theories.

But both concepts are founded on strong iterative assumptions
about mutual belief in rationality.

Backward assumption: Common belief in opponents' future
rationality regardless of what is observed.

Rationalizability: Common belief in rationality.

In some games requiring multiple levels of iterative reasoning,
these concepts have questionable descriptive or normative merit.

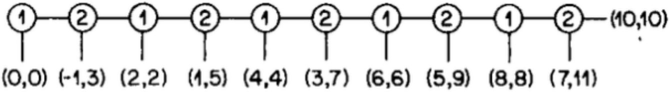
JOURNAL OF ECONOMIC THEORY **25**, 92–100 (1981)

Games of Perfect Information, Predatory Pricing and the Chain-Store Paradox

ROBERT W. ROSENTHAL

Bell Telephone Laboratories, Murray Hill, New Jersey 07974

Received April 8, 1980; revised July 2, 1980



the Centipede game

The Traveler's Dilemma (Basu, *AER P&P* 1994)

$$u_i(s_i, s_j) = \begin{cases} s_j - 2 & \text{if } s_i > s_j, \\ s_i & \text{if } s_i = s_j, \\ s_i + 2 & \text{if } s_i < s_j. \end{cases}$$

		II						
		1	2	3	4	5	6	7
	1	1, 1	3, −1	3, −1	3, −1	3, −1	3, −1	3, −1
	2	−1, 3	2, 2	4, 0	4, 0	4, 0	4, 0	4, 0
	3	−1, 3	0, 4	3, 3	5, 1	5, 1	5, 1	5, 1
I	4	−1, 3	0, 4	1, 5	4, 4	6, 2	6, 2	6, 2
	5	−1, 3	0, 4	1, 5	2, 6	5, 5	7, 3	7, 3
	6	−1, 3	0, 4	1, 5	2, 6	3, 7	6, 6	8, 4
	7	−1, 3	0, 4	1, 5	2, 6	3, 7	4, 8	7, 7

Attempts to obtain more realistic predictions in Centipede (& the TD) have been few and not entirely satisfying.

The best-known is Kreps et al. (1982), who alter the game by introducing incomplete information about opponents' preferences.

The prediction of cooperative play in this model requires

- augmenting the original game
- assuming this augmentation is common knowledge
- invoking the equilibrium knowledge assumption

Rather than using strong knowledge assumptions in a new game, we

- place the game in a population setting;
- assume agents assess actions' payoffs by direct testing
- require optimal myopic responses to these assessments

This approach generates robust cooperative behavior
in Centipede and the Traveler's Dilemma.

∴ Replacing traditional knowledge assumptions
with a simple explicit approach to assessing payoff opportunities
can provide intuitively appealing predictions of play.

Best experienced payoff dynamics

We consider population game dynamics under which agents occasionally receive opportunities to switch strategies.

A revising agent:

- chooses a set of candidate strategies according to a test-set rule τ (test-two, test-all);
- plays each strategy against κ opponents drawn at random from the opposing population, with each play of each strategy being against a newly drawn opponent;
- switches to the strategy that achieved the highest total payoff, breaking ties according to a tie-breaking rule β.

We call these best experienced payoff dynamics.

Key precursors: Osborne and Rubinstein (1998), Sethi (2000).

General form of differential equations

x, y strategy distributions for player roles 1 and 2.

Rates of change in use of each strategy are differences between inflow and outflow terms:

$$\dot{x}_i = \sum_{j \in S^1} x_j \sigma^1_{ji}(y) - x_i,$$
$$\dot{y}_i = \sum_{j \in S^2} y_j \sigma^2_{ji}(x) - y_i.$$

Conditional switch rates $\sigma^1_{ji}(y)$ and $\sigma^2_{ji}(y)$ are determined from game payoffs and the procedure described above.

Observation: *Under BEP dynamics with "stick-if-tie" tiebreaking, any pure Nash equilibrium is a rest point.*

Best experienced payoff dynamics for the Centipede game

The BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic for Centipede:

$$\dot{x}_i = \left(\sum_{k=i}^{s^2} y_k\right)\left(\sum_{m=1}^{i} y_m\right)^{s^1-i} + \sum_{k=2}^{i-1} y_k \left(\sum_{\ell=1}^{k-1} y_\ell\right)^{i-k} \left(\sum_{m=1}^{k} y_m\right)^{s^1-i} - x_i,$$

$$\dot{y}_j = \begin{cases} \left(\sum_{k=2}^{s^1} x_k\right)(x_1 + x_2)^{s^2-1} + (x_1)^{s^2} - y_1 & j = 1, \\ \left(\sum_{k=j+1}^{s^1} x_k\right)\left(\sum_{m=1}^{j+1} x_m\right)^{s^2-j} + \sum_{k=2}^{j} x_k \left(\sum_{\ell=1}^{k-1} x_\ell\right)^{j-k+1} \left(\sum_{m=1}^{k} x_m\right)^{s^2-j} - y_j & j > 1. \end{cases}$$

The BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic for Centipede:

$$\dot{x}_i = \left(\sum_{k=i}^{s^2} y_k\right)\left(\sum_{m=1}^{i} y_m\right)^{s^1-i} + \sum_{k=2}^{i-1} y_k \left(\sum_{\ell=1}^{k-1} y_\ell\right)^{i-k} \left(\sum_{m=1}^{k} y_m\right)^{s^1-i} - x_i$$

When the agent tests i, his opponent plays i or higher (so that the agent is the one who stops the game);

when the agent tests strategies above i, his opponents play i or lower.

$\Rightarrow$ Only strategy i yields the agent his highest payoff.

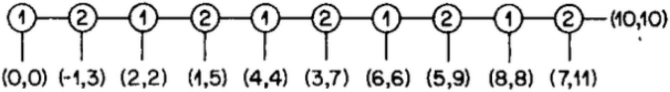
JOURNAL OF ECONOMIC THEORY **25**, 92–100 (1981)

Games of Perfect Information, Predatory Pricing and the Chain-Store Paradox

ROBERT W. ROSENTHAL

Bell Telephone Laboratories, Murray Hill, New Jersey 07974

Received April 8, 1980; revised July 2, 1980



the Centipede game

The BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic for Centipede:

$$\dot{x}_i = \left(\sum_{k=i}^{s^2} y_k\right)\left(\sum_{m=1}^{i} y_m\right)^{s^1-i} + \sum_{k=2}^{i-1} y_k \left(\sum_{\ell=1}^{k-1} y_\ell\right)^{i-k} \left(\sum_{m=1}^{k} y_m\right)^{s^1-i} - x_i$$

When the agent tests i, his opponent plays $k < i$, stopping the game;

when he tests strategies in $\{k, \ldots, i-1\}$,
his opponents play strategies less than k;

when he tests strategies above i,
his opponents play strategies k or lower.

$\Rightarrow$ Strategy i is the lowest one that achieves the optimal payoff.

Best experienced payoff dynamics for the Centipede game

The BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic for Centipede:

$$\dot{x}_i = \left(\sum_{k=i}^{s^2} y_k\right)\left(\sum_{m=1}^{i} y_m\right)^{s^1-i} + \sum_{k=2}^{i-1} y_k \left(\sum_{\ell=1}^{k-1} y_\ell\right)^{i-k} \left(\sum_{m=1}^{k} y_m\right)^{s^1-i} - x_i,$$

$$\dot{y}_j = \begin{cases} \left(\sum_{k=2}^{s^1} x_k\right)(x_1 + x_2)^{s^2-1} + (x_1)^{s^2} - y_1 & j = 1, \\ \left(\sum_{k=j+1}^{s^1} x_k\right)\left(\sum_{m=1}^{j+1} x_m\right)^{s^2-j} + \sum_{k=2}^{j} x_k \left(\sum_{\ell=1}^{k-1} x_\ell\right)^{j-k+1} \left(\sum_{m=1}^{k} x_m\right)^{s^2-j} - y_j & j > 1. \end{cases}$$

This is a system of polynomial equations with rational coefficients.

Analytical methods: computational algebra et al.

Because they are based on sampling, BEP dynamics are represented by systems of polynomial differential equations with rational coefficients.

The zeroes (= rest points) of such a system can be determined by computing a Gröbner basis for the set of polynomials.

This new set of polynomials has the same zeroes as the original set. But its zeroes can be computed by finding the roots of a single (high-degree) univariate polynomial.

Exact representations of these roots using algebraic numbers can be obtained using algorithms based on classical theorems.

From there, local stability can be assessed via linearization plus a variety of tricks (eigenvalue perturbation theorems, matrix condition number bounds. . .)

All algebraic (and numerical) computations are posted online.

Results for Centipede

We focus on test-all (later test-two), min-if-tie, $\kappa = 1$.

Proposition: *In Centipede games of lengths $d \geq 3$, the backward induction state $\xi^{\dagger}$ is repelling under the BEP$(\tau^{\text{all}}, 1, \beta^{\text{min}})$ dynamic.*

Proposition: *In Centipede games of lengths $d \geq 3$, the backward induction state $\xi^\dagger$ is repelling under the BEP$(\tau^{\text{all}}, 1, \beta^{\min})$ dynamic.*

Basic intuition:

Near the backward induction state, the most cooperative agents would obtain higher expected payoffs by stopping earlier.

However, a revising agent tests each strategy in his test set against independently drawn opponents.

He may thus test a cooperative strategy against a cooperative opponent, and less cooperative strategies against less cooperative opponents.

In this case, his best experienced payoff will come from the cooperative strategy.

Proof by a nonstandard linearization argument:
$DV(\xi^\dagger)$ has only one positive eigenvalue.

Example: the BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) dynamic in Centipede of length $d = 4$.

$$\dot{x}_1 = (y_1)^2 - x_1, \qquad \dot{y}_1 = (x_2 + x_3)(x_1 + x_2)^2 + (x_1)^3 - y_1,$$
$$\dot{x}_2 = (y_2 + y_3)(y_1 + y_2) - x_2, \qquad \dot{y}_2 = x_3 + x_2 x_1 (x_1 + x_2) - y_2,$$
$$\dot{x}_3 = y_3 + y_2 y_1 - x_3, \qquad \dot{y}_3 = x_2 (x_1)^2 + x_3 (x_1 + x_2) - y_3.$$

Linearization at $(x^\dagger, y^\dagger)$ has the positive eigenvalue 1 corresponding to eigenvector $(z^1, z^2) = ((-2, 1, 1), (-2, 1, 1))$.

$\therefore$ At $(x, y) = ((1 - 2\varepsilon, \varepsilon, \varepsilon), (1 - 2\varepsilon, \varepsilon, \varepsilon))$,
we have $(\dot{x}, \dot{y}) \approx ((-2\varepsilon, \varepsilon, \varepsilon), (-2\varepsilon, \varepsilon, \varepsilon))$.

A uniform invasion of the four unused strategies is self-reinforcing.

Example: the BEP($\tau^{\text{all}}, 1, \beta^{\min}$) dynamic in Centipede of length $d = 4$.

$$
\begin{aligned}
\dot{x}_1 &= (y_1)^2 - x_1, & \dot{y}_1 &= (x_2 + x_3)(x_1 + x_2)^2 + (x_1)^3 - y_1,\\
\dot{x}_2 &= (y_2 + y_3)(y_1 + y_2) - x_2, & \dot{y}_2 &= x_3 + x_2 x_1 (x_1 + x_2) - y_2,\\
\dot{x}_3 &= y_3 + y_2 y_1 - x_3, & \dot{y}_3 &= x_2(x_1)^2 + x_3(x_1 + x_2) - y_3.
\end{aligned}
$$

At $(x, y) = ((1 - 2\varepsilon, \varepsilon, \varepsilon), (1 - 2\varepsilon, \varepsilon, \varepsilon))$, a revising population 2 agent switches to strategy 3 (always continue) if:

i. when testing strategy 3 she meets an opponent playing 2, and when testing strategies 1 and 2 she meets opponents playing 1, or
ii. when testing strategy 3 she meets an opponent playing 3, and when testing strategy 2 she meets an opponent playing 1 or 2.

These events have total probability $\varepsilon(1 - 2\varepsilon)^2 + \varepsilon(1 - \varepsilon) \approx 2\varepsilon$.

Agents switch away from strategy 3 at rate $y_3 = \varepsilon$.

Combining the inflow and outflow terms shows that $\dot{y}_3 \approx 2\varepsilon - \varepsilon = \varepsilon$.

Proposition: *For Centipede games of lengths $3 \leq d \leq 6$, the $BEP(\tau^{\text{all}}, 1, \beta^{\text{min}})$ dynamic has exactly two rest points, $\xi^{\dagger}$ and $\xi^{*} \in \text{int}(\Xi)$. The rest point ξ^{*}, whose exact components are known, is asymptotically stable.*

Set of rest points determined using Gröbner bases.

Limitation on length of the game due to computational demands:

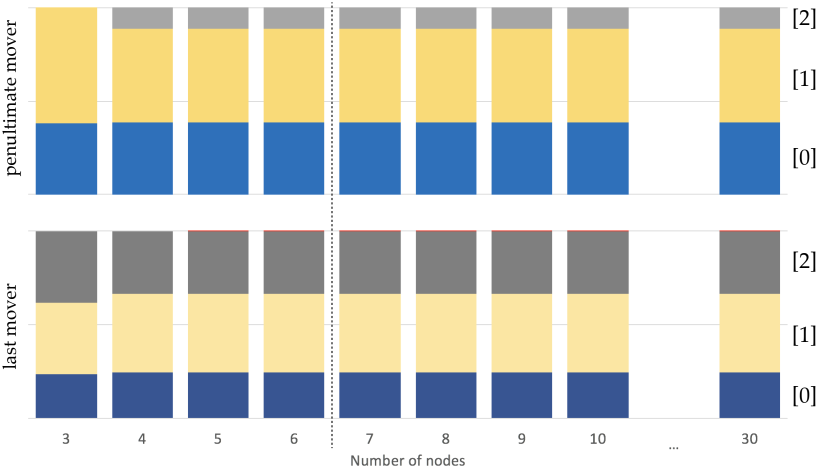
	test-all	test-two
min	6 (221)	8 (97)
stick	5 (65)	8 (128)
uniform	6 (168)	8 (128)

maximal lengths d of Centipede (degree of leading polynomial of Gröbner basis)

Local stability proved using linearization plus a variety of tricks.

Rest points and stability for longer games computed numerically.

The stable rest point in Centipede under test-all:



Each population's three most cooperative strategies predominate.

Play nearly always proceeds to one of the last five terminal nodes.

Why does each population concentrate on three strategies?

Easy to explain by examining the law of motion:

when most agents behave cooperatively, choice probabilities are almost entirely determined by a small number of terms.

$$
\begin{aligned}
\Pr(j = [0]) &\approx \quad x_{[0]}(x_{[d^1]} + \cdots + x_{[1]}) + x_{[1]}(x_{[d^1]} + \cdots + x_{[2]})^2 \\
\Pr(j = [1]) &\approx x_{[0]} + x_{[1]}(x_{[d^1]} + \cdots + x_{[2]})(x_{[d^1]} + \cdots + x_{[1]}) \\
\Pr(j = [2]) &\approx (x_{[1]} + x_{[0]})(x_{[d^1]} + \cdots + x_{[1]})^2 \\
\Pr(j = [3]) &\approx (x_{[2]} + x_{[1]} + x_{[0]})(x_{[d^1]} + \cdots + x_{[2]})^3
\end{aligned}
$$

Almost global convergence

We cannot prove analytically that $\xi^* = (x^*, y^*)$ is almost globally stable (i.e., that it attracts all solutions besides the one from $\xi^\dagger$.)

To provide strong numerical evidence of global stability, we consider the candidate Lyapunov function

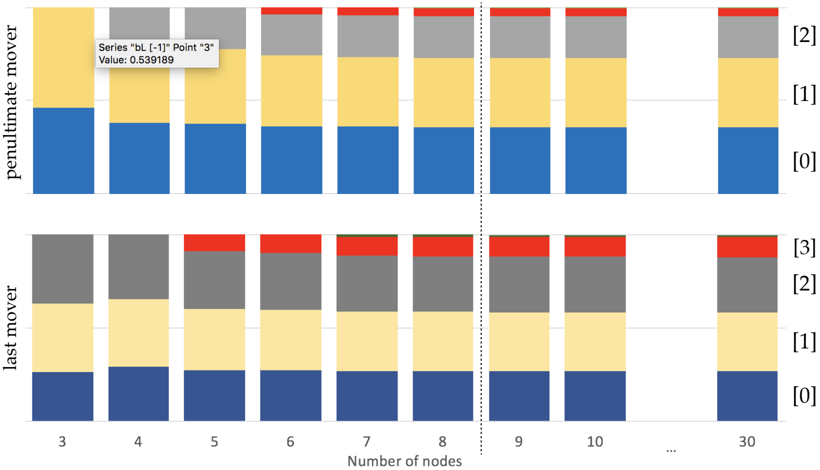
$$L(x,y) = \sum_{i=2}^{s_1}(x_i - x_i^*)^2 + \sum_{j=2}^{s_2}(y_j - y_j^*)^2$$

We numerically evaluate the time derivative $\dot{L}(x,y)$ at 10^9 points in Ξ.

For game lengths $d \leq 20$, $\dot{L}$ always evaluates to a negative number.

This is is very strong numerical evidence for almost global stability (but not a proof).

The stable rest point in Centipede under test-two:



The populations' four most cooperative strategies predominate.

Play nearly always proceeds to one of the last seven terminal nodes.

Results for the Traveler's Dilemma

$$u_i(s_i, s_j) = \begin{cases} s_j - 2 & \text{if } s_i > s_j, \\ s_i & \text{if } s_i = s_j, \\ s_i + 2 & \text{if } s_i < s_j. \end{cases}$$

		II						
		1	2	3	4	5	6	7
	1	1, 1	3, −1	3, −1	3, −1	3, −1	3, −1	3, −1
	2	−1, 3	2, 2	4, 0	4, 0	4, 0	4, 0	4, 0
	3	−1, 3	0, 4	3, 3	5, 1	5, 1	5, 1	5, 1
I	4	−1, 3	0, 4	1, 5	4, 4	6, 2	6, 2	6, 2
	5	−1, 3	0, 4	1, 5	2, 6	5, 5	7, 3	7, 3
	6	−1, 3	0, 4	1, 5	2, 6	3, 7	6, 6	8, 4
	7	−1, 3	0, 4	1, 5	2, 6	3, 7	4, 8	7, 7

Because the game is symmetric, we look at a single-population model.

We again focus on test-all/test-two, min-if-tie, $\kappa = 1$.

Results for test-all:

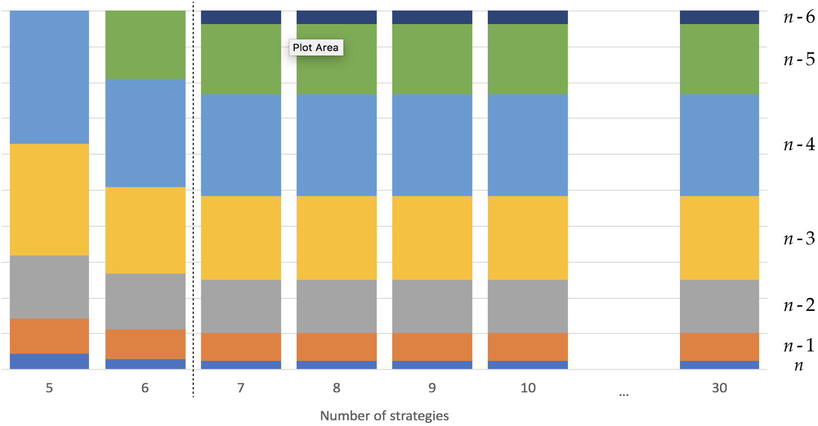
Proposition: *In Traveler's Dilemma games with* $n \geq 5$ *strategies, the rationalizable state* $x^\dagger$ *is unstable under the BEP*$(\tau^{\text{all}}, 1, \beta^{\text{min}})$ *dynamic.*

Proposition: *In Traveler's Dilemma games with* $5 \leq n \leq 6$ *strategies, the BEP*$(\tau^{\text{all}}, 1, \beta^{\text{min}})$ *dynamic has exactly two rest points,* $x^\dagger$ *and* $x^* \in \text{int}(\Xi)$*. The rest point* x^**, whose exact components are known, is asymptotically stable.*

The stable rest point x^* in the Traveler's Dilemma under test-all:

Weights are essentially independent of $n \geq 7$.

Play at x^* is concentrated on the seven highest strategies.



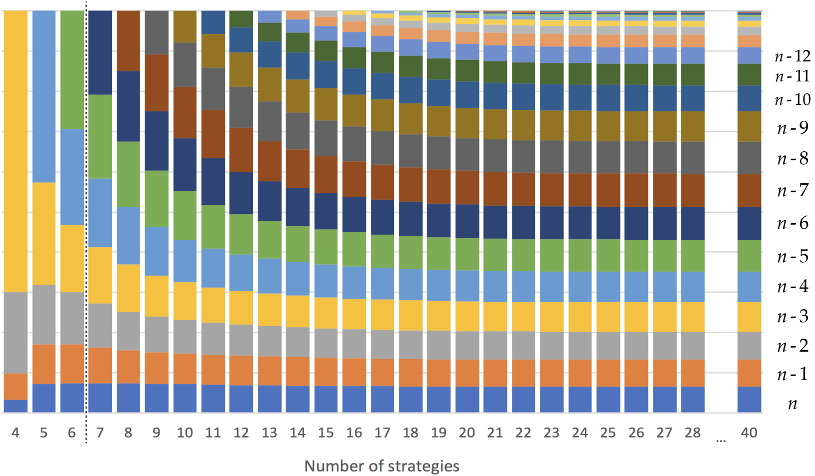
Results for test-two:

Proposition: *In Traveler's Dilemma games with* $n \geq 7$ *strategies, the rationalizable state* $x^\dagger$ *is unstable under the BEP*$(\tau^{\text{two}}, 1, \beta^{\text{min}})$ *dynamic.*

Proposition: *In Traveler's Dilemma games with* $4 \leq n \leq 6$ *strategies, the BEP*$(\tau^{\text{two}}, 1, \beta^{\text{min}})$ *dynamic has exactly two rest points,* $x^\dagger$ *and* $x^* \in \text{int}(\Xi)$. *The rest point* x^*, *whose exact components are known, is asymptotically stable.*

The stable rest point x^* in the Traveler's Dilemma under test-two:

Play is concentrated on the ≈ 20 highest strategies.



Robustness (in Centipede)

Robustness to including "backward induction agents"

What happens if we introduce heterogeneity, with some agents who always stop immediately?

$d = 4,\ \tau = \tau^{\mathrm{all}}$

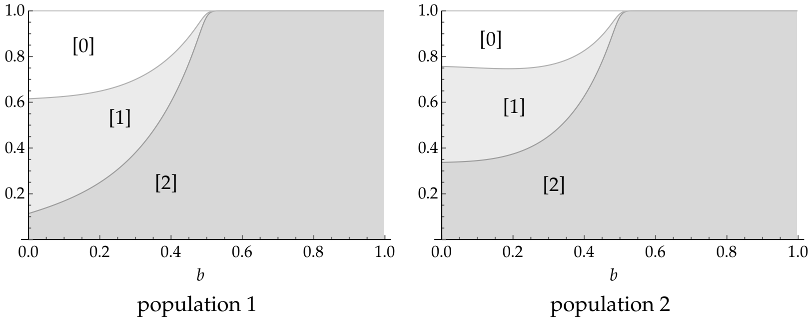


Figure: Behavior of BEP(τ^{all}, 1, $\beta^{\min}$) agents at the stable rest point when proportion $b \in [0, 1]$ always stops immediately ($d = 4$).

$d = 20,\ \tau = \tau^{\text{all}}$

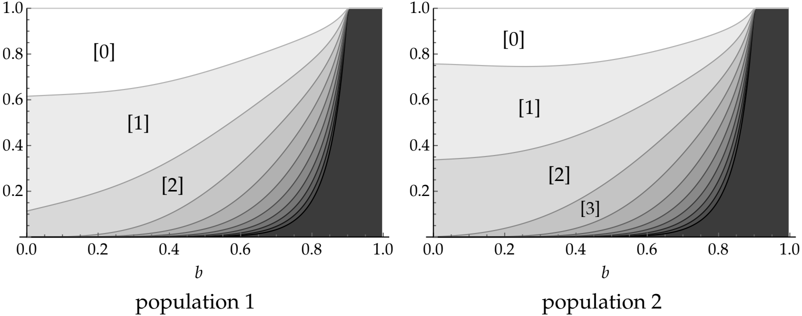


Figure: Behavior of BEP($\tau^{\text{all}}, 1, \beta^{\text{min}}$) agents at the stable rest point when proportion $b \in [0, 1]$ always stops immediately ($d = 20$).

$d = 20, \ \tau = \tau^{\text{two}}$

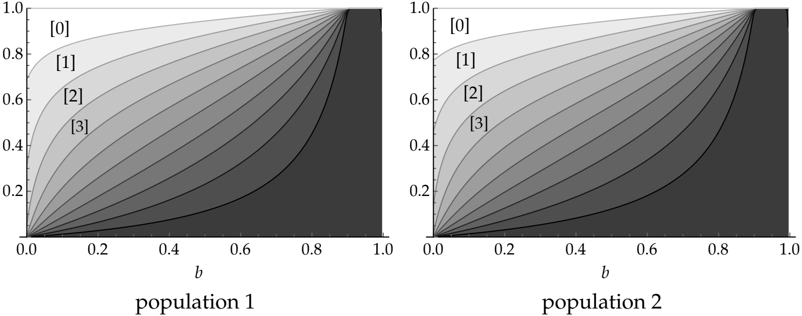


Figure: Behavior of BEP($\tau^{\text{two}}, 1, \beta^{\text{min}}$) agents at the stable rest point when proportion $b \in [0, 1]$ always stops immediately ($d = 20$).

Larger numbers of trials

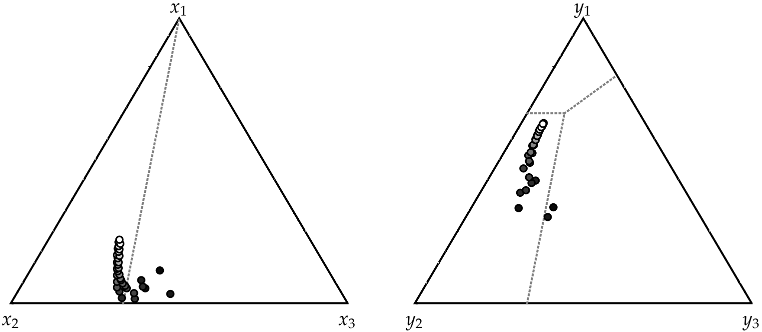
So far we have assumed that agents test the strategies in their test sets exactly once.

What happens if we increase the number of trials κ of each tested strategy?

The case $\kappa = \infty$ corresponds to the best response dynamic.

Results of Xu (2016) imply that every solution converges to the set of Nash equilibria $\Rightarrow$ all population 1 agents stop immediately.

What happens in intermediate cases?



$\kappa = 1, \ldots, 34$

Figure: The stable rest point in Centipede of length $d = 4$ under BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamics for $\kappa = 1, \ldots, 34$ trials of each tested strategy. Lighter shading corresponds to larger numbers of trials.

Intuition from central limit theorem.

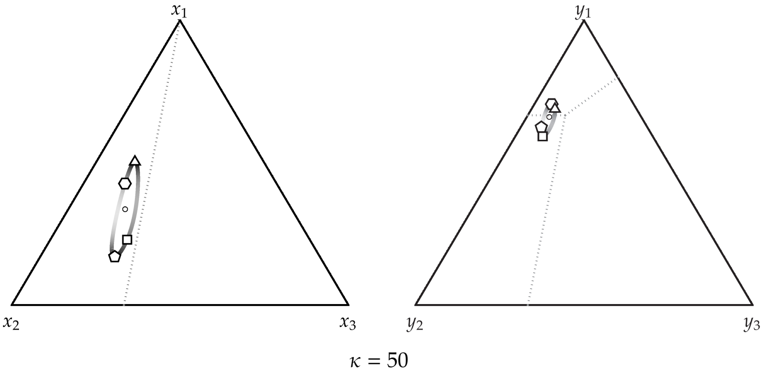


Figure: Stable cycle in Centipede of length $d = 4$ under BEP$(\tau^{\text{all}}, \kappa, \beta^{\min})$ dynamics for $\kappa = 50$. Lighter shading represents faster motion.

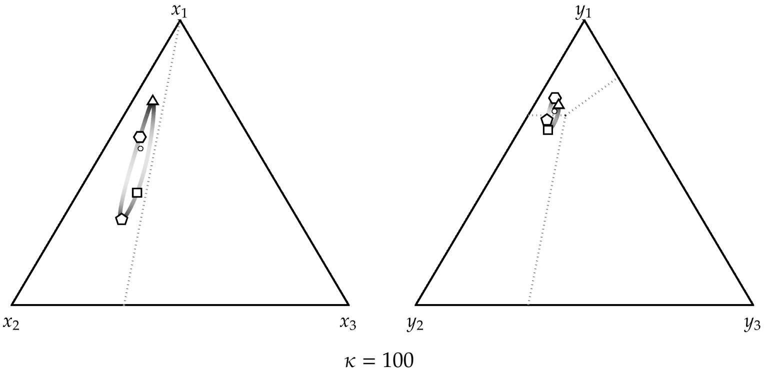


Figure: Stable cycle in Centipede of length $d = 4$ under BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamics for $\kappa = 100$. Lighter shading represents faster motion.

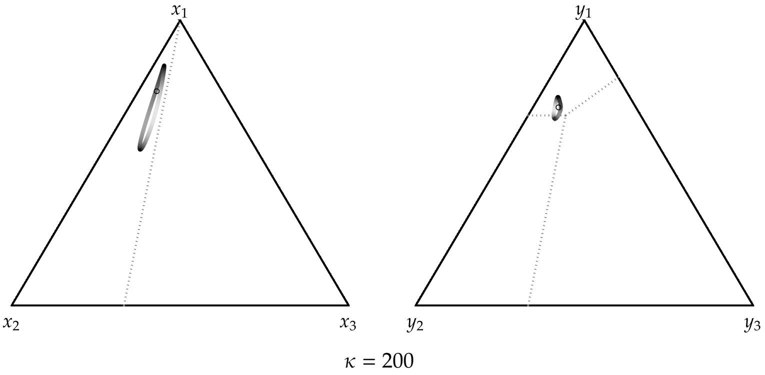


Figure: Stable cycle in Centipede of length $d = 4$ under BEP($\tau^{\text{all}}, \kappa, \beta^{\text{min}}$) dynamics for $\kappa = 200$. Lighter shading represents faster motion.

Conclusion

Best experienced payoff dynamics do not assume that agents correctly anticipate others behavior, but do assume that agents choose optimally given the information they possess.

In Centipede and the Traveler's Dilemma, these dynamics lead to stable and robust cooperative behavior.

Elementary knowledge assumptions can lead to intuitive predictions in games where traditional iterative solution concepts do not.

Methods from computational algebra etc. should prove useful for studying other game models that rely on sampling and other polynomial procedures.