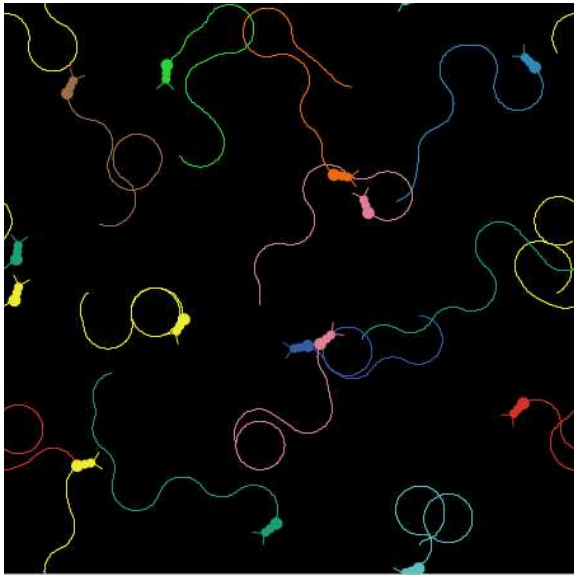
Multi-agent learning

Fictitious Play

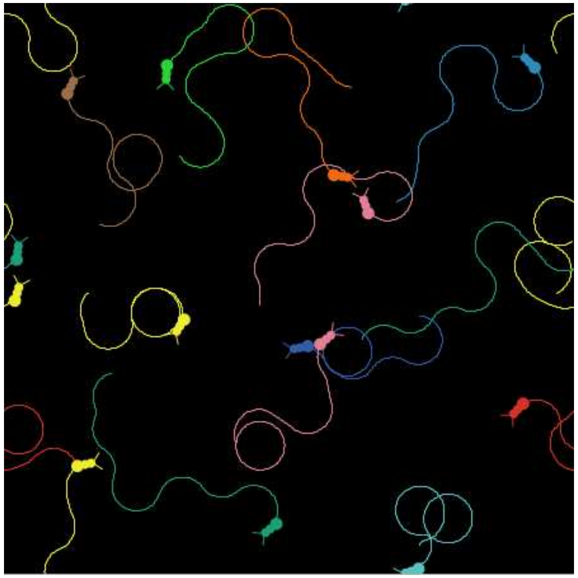
Gerard Vreeswijk, Intelligent Software Systems, Computer Science Department, Faculty of Sciences, Utrecht University, The Netherlands.

Wednesday 13th May, 2020

Fictitious play: motivation

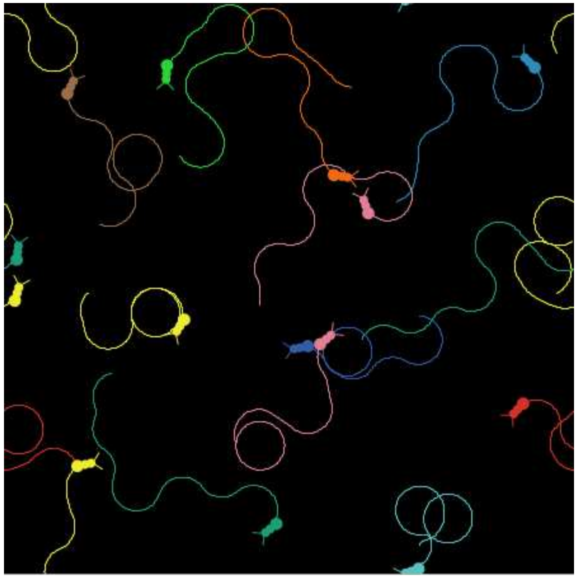


Fictitious play: motivation



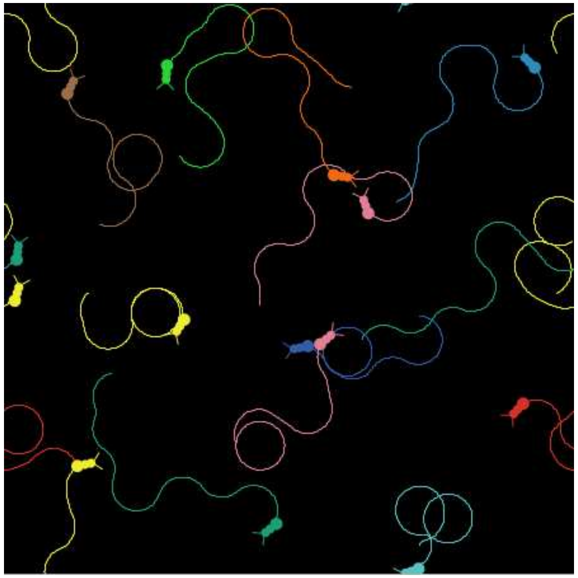
- One of the most important, if not the most important,

Fictitious play: motivation



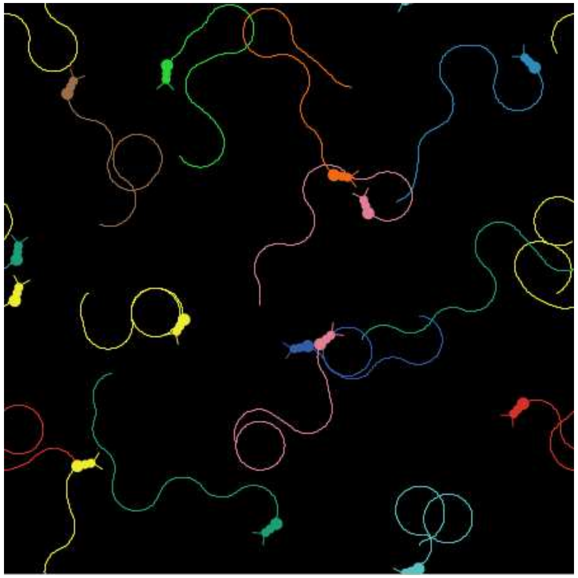
- One of the most important, if not the most important, representative of a follower strategy.

Fictitious play: motivation



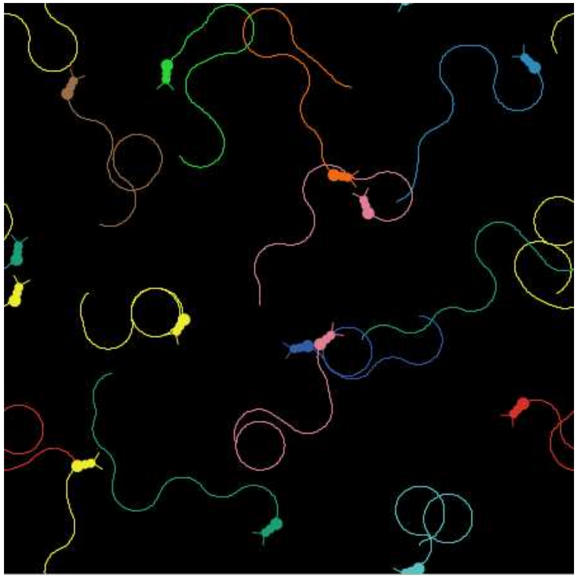
- One of the most important, if not the most important, representative of a follower strategy.
- Rather than considering your own payoffs, monitor the behaviour of your opponent(s).

Fictitious play: motivation



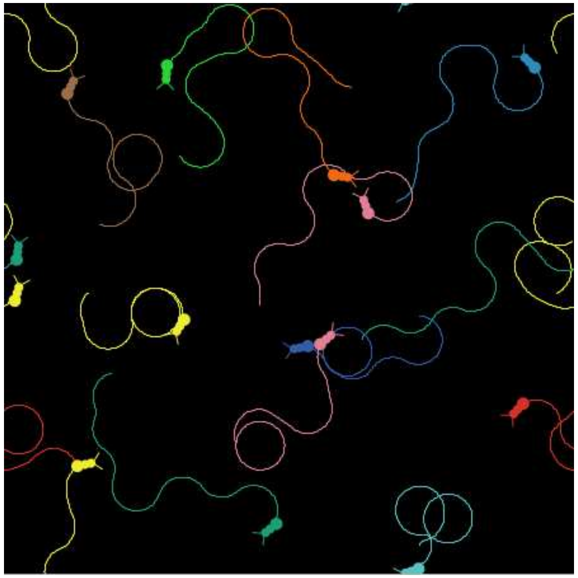
- One of the most important, if not the most important, representative of a follower strategy.
- Rather than considering your own payoffs, monitor the behaviour of your opponent(s).
- Project the behaviour of an opponent onto a single mixed strategy, s.

Fictitious play: motivation



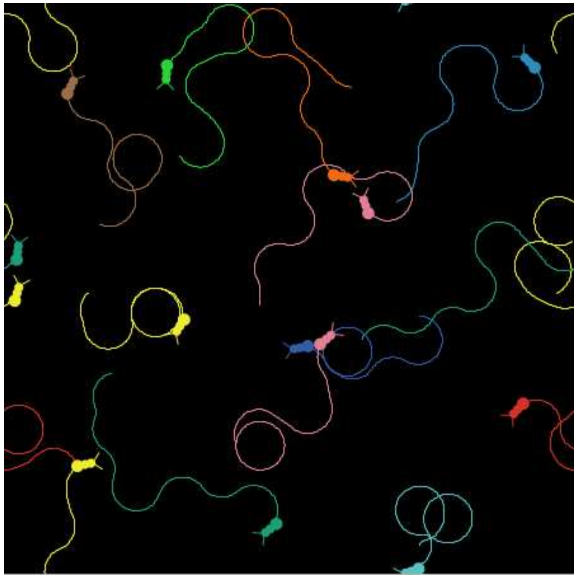
- One of the most important, if not the most important, representative of a follower strategy.
- Rather than considering your own payoffs, monitor the behaviour of your opponent(s).
- Project the behaviour of an opponent onto a single mixed strategy, s.
- Then issue a best response to s.

Fictitious play: motivation



- One of the most important, if not the most important, representative of a follower strategy.
- Rather than considering your own payoffs, monitor the behaviour of your opponent(s).
- Project the behaviour of an opponent onto a single mixed strategy, s.
- Then issue a best response to s.
- Brown (1951): explanation for Nash equilibrium play.

Fictitious play: motivation



- One of the most important, if not the most important, representative of a follower strategy.
- Rather than considering your own payoffs, monitor the behaviour of your opponent(s).
- Project the behaviour of an opponent onto a single mixed strategy, s.
- Then issue a best response to s.
- Brown (1951): explanation for Nash equilibrium play. In terms of current use, the name actually is a bit of a misnomer, since play actually occurs (Berger, 2005).

Plan for today

Plan for today

Part I. Best reply strategy

Plan for today

Part I. Best reply strategy

1. Pure fictitious play.

Plan for today

Part I. Best reply strategy

1. Pure fictitious play.
2. Results that connect pure fictitious play to Nash equilibria.

Plan for today

Part I. Best reply strategy

1. Pure fictitious play.
2. Results that connect pure fictitious play to Nash equilibria.

Part II. Extensions and approximations of fictitious play

Plan for today

Part I. Best reply strategy

1. Pure fictitious play.

2. Results that connect pure fictitious play to Nash equilibria.

Part II. Extensions and approximations of fictitious play

1. Smoothed fictitious play.

Plan for today

Part I. Best reply strategy

1. Pure fictitious play.
2. Results that connect pure fictitious play to Nash equilibria.

Part II. Extensions and approximations of fictitious play

1. Smoothed fictitious play.
2. Exponential regret matching.

Plan for today

Part I. Best reply strategy

1. Pure fictitious play.
2. Results that connect pure fictitious play to Nash equilibria.

Part II. Extensions and approximations of fictitious play

1. Smoothed fictitious play.
2. Exponential regret matching.
3. No-regret property of smoothed fictitious play (Fudenberg *et al.*, 1995).

Plan for today

Part I. Best reply strategy

1. Pure fictitious play.
2. Results that connect pure fictitious play to Nash equilibria.

Part II. Extensions and approximations of fictitious play

1. Smoothed fictitious play.
2. Exponential regret matching.
3. No-regret property of smoothed fictitious play (Fudenberg *et al.*, 1995).
4. Convergence when players have limited resources (Young, 1998).

Plan for today

Part I. Best reply strategy

1. Pure fictitious play.
2. Results that connect pure fictitious play to Nash equilibria.

Part II. Extensions and approximations of fictitious play

1. Smoothed fictitious play.
2. Exponential regret matching.
3. No-regret property of smoothed fictitious play (Fudenberg *et al.*, 1995).
4. Convergence when players have limited resources (Young, 1998).

Shoham *et al.* (2009): *Multi-agent Systems*. Ch. 7: "Learning and Teaching". H. Young (2004): *Strategic Learning and it Limits*, Oxford UP. D. Fudenberg and D.K. Levine (1998), *The Theory of Learning in Games*, MIT Press.

Part I:
Pure fictitious play

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0, 0)$, $(0.5, 0.5)$, and $(1, 1)$.

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Round	A's action	B's action	A's beliefs	B's beliefs
0.				

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0, 0)$, $(0.5, 0.5)$, and $(1, 1)$.

* = chosen randomly.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0, 0)$	$(0, 0)$

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0,0)$	$(0,0)$
1.	L*	R*		

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0, 0)$, $(0.5, 0.5)$, and $(1, 1)$.

* = chosen randomly.

Round	**A's action**	**B's action**	**A's beliefs**	**B's beliefs**
0.			$(0, 0)$	$(0, 0)$
1.	L*	R*	$(0, 1)$	$(1, 0)$

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0, 0)$, $(0.5, 0.5)$, and $(1, 1)$.

* = chosen randomly.

Round	***A*'s action**	***B*'s action**	***A*'s beliefs**	***B*'s beliefs**
0.			$(0, 0)$	$(0, 0)$
1.	L*	R*	$(0, 1)$	$(1, 0)$
2.	R	L		

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0,0)$	$(0,0)$
1.	L*	R*	$(0,1)$	$(1,0)$
2.	R	L	$(1,1)$	$(1,1)$

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0, 0)$, $(0.5, 0.5)$, and $(1, 1)$.

* = chosen randomly.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0, 0)$	$(0, 0)$
1.	L*	R*	$(0, 1)$	$(1, 0)$
2.	R	L	$(1, 1)$	$(1, 1)$
3.	L*	R*		

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Round	**A's action**	**B's action**	**A's beliefs**	**B's beliefs**
0.			$(0,0)$	$(0,0)$
1.	L*	R*	$(0,1)$	$(1,0)$
2.	R	L	$(1,1)$	$(1,1)$
3.	L*	R*	$(1,2)$	$(2,1)$

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Round	**A's action**	**B's action**	**A's beliefs**	**B's beliefs**
0.			$(0,0)$	$(0,0)$
1.	L*	R*	$(0,1)$	$(1,0)$
2.	R	L	$(1,1)$	$(1,1)$
3.	L*	R*	$(1,2)$	$(2,1)$
4.	R	L		

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0,0)$	$(0,0)$
1.	L*	R*	$(0,1)$	$(1,0)$
2.	R	L	$(1,1)$	$(1,1)$
3.	L*	R*	$(1,2)$	$(2,1)$
4.	R	L	$(2,2)$	$(2,2)$

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0,0)$	$(0,0)$
1.	L*	R*	$(0,1)$	$(1,0)$
2.	R	L	$(1,1)$	$(1,1)$
3.	L*	R*	$(1,2)$	$(2,1)$
4.	R	L	$(2,2)$	$(2,2)$
5.	R*	R*		

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0,0)$	$(0,0)$
1.	L*	R*	$(0,1)$	$(1,0)$
2.	R	L	$(1,1)$	$(1,1)$
3.	L*	R*	$(1,2)$	$(2,1)$
4.	R	L	$(2,2)$	$(2,2)$
5.	R*	R*	$(2,3)$	$(2,3)$

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0, 0)$, $(0.5, 0.5)$, and $(1, 1)$.

* = chosen randomly.

Round	***A*'s action**	***B*'s action**	***A*'s beliefs**	***B*'s beliefs**
0.			$(0, 0)$	$(0, 0)$
1.	L*	R*	$(0, 1)$	$(1, 0)$
2.	R	L	$(1, 1)$	$(1, 1)$
3.	L*	R*	$(1, 2)$	$(2, 1)$
4.	R	L	$(2, 2)$	$(2, 2)$
5.	R*	R*	$(2, 3)$	$(2, 3)$
6.	R	R		

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0,0)$	$(0,0)$
1.	L*	R*	$(0,1)$	$(1,0)$
2.	R	L	$(1,1)$	$(1,1)$
3.	L*	R*	$(1,2)$	$(2,1)$
4.	R	L	$(2,2)$	$(2,2)$
5.	R*	R*	$(2,3)$	$(2,3)$
6.	R	R	$(2,4)$	$(2,4)$

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0,0)$, $(0.5,0.5)$, and $(1,1)$.

* = chosen randomly.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0,0)$	$(0,0)$
1.	L*	R*	$(0,1)$	$(1,0)$
2.	R	L	$(1,1)$	$(1,1)$
3.	L*	R*	$(1,2)$	$(2,1)$
4.	R	L	$(2,2)$	$(2,2)$
5.	R*	R*	$(2,3)$	$(2,3)$
6.	R	R	$(2,4)$	$(2,4)$
7.	R	R		

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0, 0)$, $(0.5, 0.5)$, and $(1, 1)$.

* = chosen randomly.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0, 0)$	$(0, 0)$
1.	L*	R*	$(0, 1)$	$(1, 0)$
2.	R	L	$(1, 1)$	$(1, 1)$
3.	L*	R*	$(1, 2)$	$(2, 1)$
4.	R	L	$(2, 2)$	$(2, 2)$
5.	R*	R*	$(2, 3)$	$(2, 3)$
6.	R	R	$(2, 4)$	$(2, 4)$
7.	R	R	$(2, 5)$	$(2, 5)$

Repeated coordination game

Players receive a positive payoff iff they coordinate. This game possesses three Nash equilibria, viz. $(0, 0)$, $(0.5, 0.5)$, and $(1, 1)$.

* = chosen randomly.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0, 0)$	$(0, 0)$
1.	L*	R*	$(0, 1)$	$(1, 0)$
2.	R	L	$(1, 1)$	$(1, 1)$
3.	L*	R*	$(1, 2)$	$(2, 1)$
4.	R	L	$(2, 2)$	$(2, 2)$
5.	R*	R*	$(2, 3)$	$(2, 3)$
6.	R	R	$(2, 4)$	$(2, 4)$
7.	R	R	$(2, 5)$	$(2, 5)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Some types of Nash equilibria

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation, i.e. $B(s) = \{s\}$.

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation, i.e. $B(s) = \{s\}$. Opposite: weak Nash.

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation, i.e. $B(s) = \{s\}$. Opposite: weak Nash.
- Pure Nash. Every party maintains a pure strategy.

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation, i.e. $B(s) = \{s\}$. Opposite: weak Nash.
- Pure Nash. Every party maintains a pure strategy. Opposite: mixed (one, more, all may be pure), and fully mixed (all mixed).

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation, i.e. $B(s) = \{s\}$. Opposite: weak Nash.
- Pure Nash. Every party maintains a pure strategy. Opposite: mixed (one, more, all may be pure), and fully mixed (all mixed).
- Strong Nash. No coalition wins by unilateral deviation.

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation, i.e. $B(s) = \{s\}$. Opposite: weak Nash.
- Pure Nash. Every party maintains a pure strategy. Opposite: mixed (one, more, all may be pure), and fully mixed (all mixed).
- Strong Nash. No coalition wins by unilateral deviation. Implies Pareto-optimality.

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation, i.e. $B(s) = \{s\}$. Opposite: weak Nash.
- Pure Nash. Every party maintains a pure strategy. Opposite: mixed (one, more, all may be pure), and fully mixed (all mixed).
- Strong Nash. No coalition wins by unilateral deviation. Implies Pareto-optimality. Opposite: ? (Not weak: a Nash equilibrium can be both strong and weak, either, or neither.)

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation, i.e. $B(s) = \{s\}$. Opposite: weak Nash.
- Pure Nash. Every party maintains a pure strategy. Opposite: mixed (one, more, all may be pure), and fully mixed (all mixed).
- Strong Nash. No coalition wins by unilateral deviation. Implies Pareto-optimality. Opposite: ? (Not weak: a Nash equilibrium can be both strong and weak, either, or neither.)
- Stable Nash. No party wins by small unilateral deviation, moreover the the one who deviates then loses.

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation, i.e. $B(s) = \{s\}$. Opposite: weak Nash.
- Pure Nash. Every party maintains a pure strategy. Opposite: mixed (one, more, all may be pure), and fully mixed (all mixed).
- Strong Nash. No coalition wins by unilateral deviation. Implies Pareto-optimality. Opposite: ? (Not weak: a Nash equilibrium can be both strong and weak, either, or neither.)
- Stable Nash. No party wins by small unilateral deviation, moreover the the one who deviates then loses. Opposite: unstable.

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation, i.e. $B(s) = \{s\}$. Opposite: weak Nash.
- Pure Nash. Every party maintains a pure strategy. Opposite: mixed (one, more, all may be pure), and fully mixed (all mixed).
- Strong Nash. No coalition wins by unilateral deviation. Implies Pareto-optimality. Opposite: ? (Not weak: a Nash equilibrium can be both strong and weak, either, or neither.)
- Stable Nash. No party wins by small unilateral deviation, moreover the the one who deviates then loses. Opposite: unstable.

There are many more equilibrium types (semi-strict, locally stable, ϵ-equilibrium, correlated, coarse correlated, ...).

Some types of Nash equilibria

- Nash. No party wins by unilateral deviation.
- Strict Nash. Every party loses by unilateral deviation, i.e. $B(s) = \{s\}$. Opposite: weak Nash.
- Pure Nash. Every party maintains a pure strategy. Opposite: mixed (one, more, all may be pure), and fully mixed (all mixed).
- Strong Nash. No coalition wins by unilateral deviation. Implies Pareto-optimality. Opposite: ? (Not weak: a Nash equilibrium can be both strong and weak, either, or neither.)
- Stable Nash. No party wins by small unilateral deviation, moreover the the one who deviates then loses. Opposite: unstable.

There are many more equilibrium types (semi-strict, locally stable, ϵ-equilibrium, correlated, coarse correlated, ...). Which equilibrium types imply which other equilibrium types is an interesting question.

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary)

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary) if it is the case that whenever S is in state s^*, it remains so.

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary) if it is the case that whenever S is in state s^*, it remains so.

Theorem. *Suppose a pure strategy profile is a steady state of fictitious play.*

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary) if it is the case that whenever S is in state s^*, it remains so.

Theorem. *Suppose a pure strategy profile is a steady state of fictitious play. Then it must be a (possibly weak) pure Nash equilibrium in the stage game.*

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary) if it is the case that whenever S is in state s^*, it remains so.

Theorem. *Suppose a pure strategy profile is a steady state of fictitious play. Then it must be a (possibly weak) pure Nash equilibrium in the stage game.*

Proof.

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary) if it is the case that whenever S is in state s^*, it remains so.

Theorem. *Suppose a pure strategy profile is a steady state of fictitious play. Then it must be a (possibly weak) pure Nash equilibrium in the stage game.*

Proof. Suppose $a = (a^1, \ldots, a^n)$ is a steady state.

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary) if it is the case that whenever S is in state s^*, it remains so.

Theorem. *Suppose a pure strategy profile is a steady state of fictitious play. Then it must be a (possibly weak) pure Nash equilibrium in the stage game.*

Proof. Suppose $a = (a^1, \dots, a^n)$ is a steady state. Let $1 \leq i \leq n$ be the index of an arbitrary player.

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary) if it is the case that whenever S is in state s^*, it remains so.

Theorem. *Suppose a pure strategy profile is a steady state of fictitious play. Then it must be a (possibly weak) pure Nash equilibrium in the stage game.*

Proof. Suppose $a = (a^1, \ldots, a^n)$ is a steady state. Let $1 \leq i \leq n$ be the index of an arbitrary player. At every round, i's opponent model is a^{-i}.

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary) if it is the case that whenever S is in state s^*, it remains so.

Theorem. *Suppose a pure strategy profile is a steady state of fictitious play. Then it must be a (possibly weak) pure Nash equilibrium in the stage game.*

Proof. Suppose $a = (a^1, \ldots, a^n)$ is a steady state. Let $1 \leq i \leq n$ be the index of an arbitrary player. At every round, i's opponent model is a^{-i}. By definition of fictitious play, player i must have played a best response

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary) if it is the case that whenever S is in state s^*, it remains so.

Theorem. *Suppose a pure strategy profile is a steady state of fictitious play. Then it must be a (possibly weak) pure Nash equilibrium in the stage game.*

Proof. Suppose $a = (a^1, \dots, a^n)$ is a steady state. Let $1 \leq i \leq n$ be the index of an arbitrary player. At every round, i's opponent model is a^{-i}. By definition of fictitious play, player i must have played a best response, i.e.,

$$a^i \in \mathrm{BR}(a^{-i}).$$

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary) if it is the case that whenever S is in state s^*, it remains so.

Theorem. *Suppose a pure strategy profile is a steady state of fictitious play. Then it must be a (possibly weak) pure Nash equilibrium in the stage game.*

Proof. Suppose $a = (a^1, \dots, a^n)$ is a steady state. Let $1 \leq i \leq n$ be the index of an arbitrary player. At every round, i's opponent model is a^{-i}. By definition of fictitious play, player i must have played a best response, i.e.,

$$a^i \in \mathrm{BR}(a^{-i}).$$

Since i was arbitrary, this holds for every player i.

Steady states are pure (but possibly weak) NE

Definition (Steady state). Let S by a dynamic process. A state s^* is said to be steady (or absorbing, or stationary) if it is the case that whenever S is in state s^*, it remains so.

Theorem. *Suppose a pure strategy profile is a steady state of fictitious play. Then it must be a (possibly weak) pure Nash equilibrium in the stage game.*

Proof. Suppose $a = (a^1, \dots, a^n)$ is a steady state. Let $1 \leq i \leq n$ be the index of an arbitrary player. At every round, i's opponent model is a^{-i}. By definition of fictitious play, player i must have played a best response, i.e.,

$$a^i \in \mathrm{BR}(a^{-i}).$$

Since i was arbitrary, this holds for every player i. The action profile a is therefore a Nash equilibrium. $\square$

Pure strict Nash equilibria are steady states

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game.*

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Notice the use of terminology:

■

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Notice the use of terminology:

- "pure strategy profile" for Nash equilibria.
-

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Notice the use of terminology:

- "pure strategy profile" for Nash equilibria.
- "steady state" for fictitious play.

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Notice the use of terminology:

- "pure strategy profile" for Nash equilibria.
- "steady state" for fictitious play.

Proof.

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a <u>strict</u> Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Notice the use of terminology:

- "pure strategy profile" for Nash equilibria.
- "steady state" for fictitious play.

Proof. Suppose the pure action profile a is a strict Nash equilibrium.

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Notice the use of terminology:

- "pure strategy profile" for Nash equilibria.
- "steady state" for fictitious play.

Proof. Suppose the pure action profile a is a strict Nash equilibrium. Suppose a is played at round t.

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Notice the use of terminology:

- "pure strategy profile" for Nash equilibria.
- "steady state" for fictitious play.

Proof. Suppose the pure action profile a is a strict Nash equilibrium. Suppose a is played at round t. Because a is strict, a_i is the unique best response to a_{-i}, for each i.

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Notice the use of terminology:

- "pure strategy profile" for Nash equilibria.
- "steady state" for fictitious play.

Proof. Suppose the pure action profile a is a strict Nash equilibrium. Suppose a is played at round t. Because a is strict, a_i is the unique best response to a_{-i}, for each i. Therefore, the action profile a will be played in round $t+1$ again. $\square$

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Notice the use of terminology:

- "pure strategy profile" for Nash equilibria.
- "steady state" for fictitious play.

Proof. Suppose the pure action profile a is a strict Nash equilibrium. Suppose a is played at round t. Because a is strict, a_i is the unique best response to a_{-i}, for each i. Therefore, the action profile a will be played in round $t+1$ again. $\square$

Summary of the two theorems:

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Notice the use of terminology:

- "pure strategy profile" for Nash equilibria.
- "steady state" for fictitious play.

Proof. Suppose the pure action profile a is a strict Nash equilibrium. Suppose a is played at round t. Because a is strict, a_i is the unique best response to a_{-i}, for each i. Therefore, the action profile a will be played in round $t+1$ again. $\square$

Summary of the two theorems:

$$\text{Strict and pure Nash} \quad \Rightarrow \quad \text{Steady state} \quad \Rightarrow \quad \text{Pure Nash.}$$

Pure strict Nash equilibria are steady states

Theorem. *Suppose a pure strategy profile is a strict Nash equilibrium of a stage game. Then it must be a steady state of fictitious play.*

Notice the use of terminology:

- "pure strategy profile" for Nash equilibria.
- "steady state" for fictitious play.

Proof. Suppose the pure action profile a is a strict Nash equilibrium. Suppose a is played at round t. Because a is strict, a_i is the unique best response to a_{-i}, for each i. Therefore, the action profile a will be played in round $t + 1$ again. $\square$

Summary of the two theorems:

Strict and pure Nash $\Rightarrow$ Steady state $\Rightarrow$ Pure Nash.

But what if all equilibria are mixed?

What if all equilibria are mixed?

Example. Matching Pennies.

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game.

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A's action	B's action	A's beliefs	B's beliefs
0.				

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(1.5, 2.0)$	$(2.0, 1.5)$

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T		

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H		

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H		

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$
4.	H	H		

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	**A's action**	**B's action**	**A's beliefs**	**B's beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$
4.	H	H	$(4.5, 3.0)$	$(3.0, 4.5)$

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$
4.	H	H	$(4.5, 3.0)$	$(3.0, 4.5)$
5.	H	H		

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$
4.	H	H	$(4.5, 3.0)$	$(3.0, 4.5)$
5.	H	H	$(5.5, 3.0)$	$(4.0, 4.5)$

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$
4.	H	H	$(4.5, 3.0)$	$(3.0, 4.5)$
5.	H	H	$(5.5, 3.0)$	$(4.0, 4.5)$
6.	H	H		

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$
4.	H	H	$(4.5, 3.0)$	$(3.0, 4.5)$
5.	H	H	$(5.5, 3.0)$	$(4.0, 4.5)$
6.	H	H	$(6.5, 3.0)$	$(5.0, 4.5)$

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$
4.	H	H	$(4.5, 3.0)$	$(3.0, 4.5)$
5.	H	H	$(5.5, 3.0)$	$(4.0, 4.5)$
6.	H	H	$(6.5, 3.0)$	$(5.0, 4.5)$
7.	H	T		

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	**A's action**	**B's action**	**A's beliefs**	**B's beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$
4.	H	H	$(4.5, 3.0)$	$(3.0, 4.5)$
5.	H	H	$(5.5, 3.0)$	$(4.0, 4.5)$
6.	H	H	$(6.5, 3.0)$	$(5.0, 4.5)$
7.	H	T	$(6.5, 4.0)$	$(6.0, 4.5)$

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	**A's action**	**B's action**	**A's beliefs**	**B's beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$
4.	H	H	$(4.5, 3.0)$	$(3.0, 4.5)$
5.	H	H	$(5.5, 3.0)$	$(4.0, 4.5)$
6.	H	H	$(6.5, 3.0)$	$(5.0, 4.5)$
7.	H	T	$(6.5, 4.0)$	$(6.0, 4.5)$
8.	H	T		

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

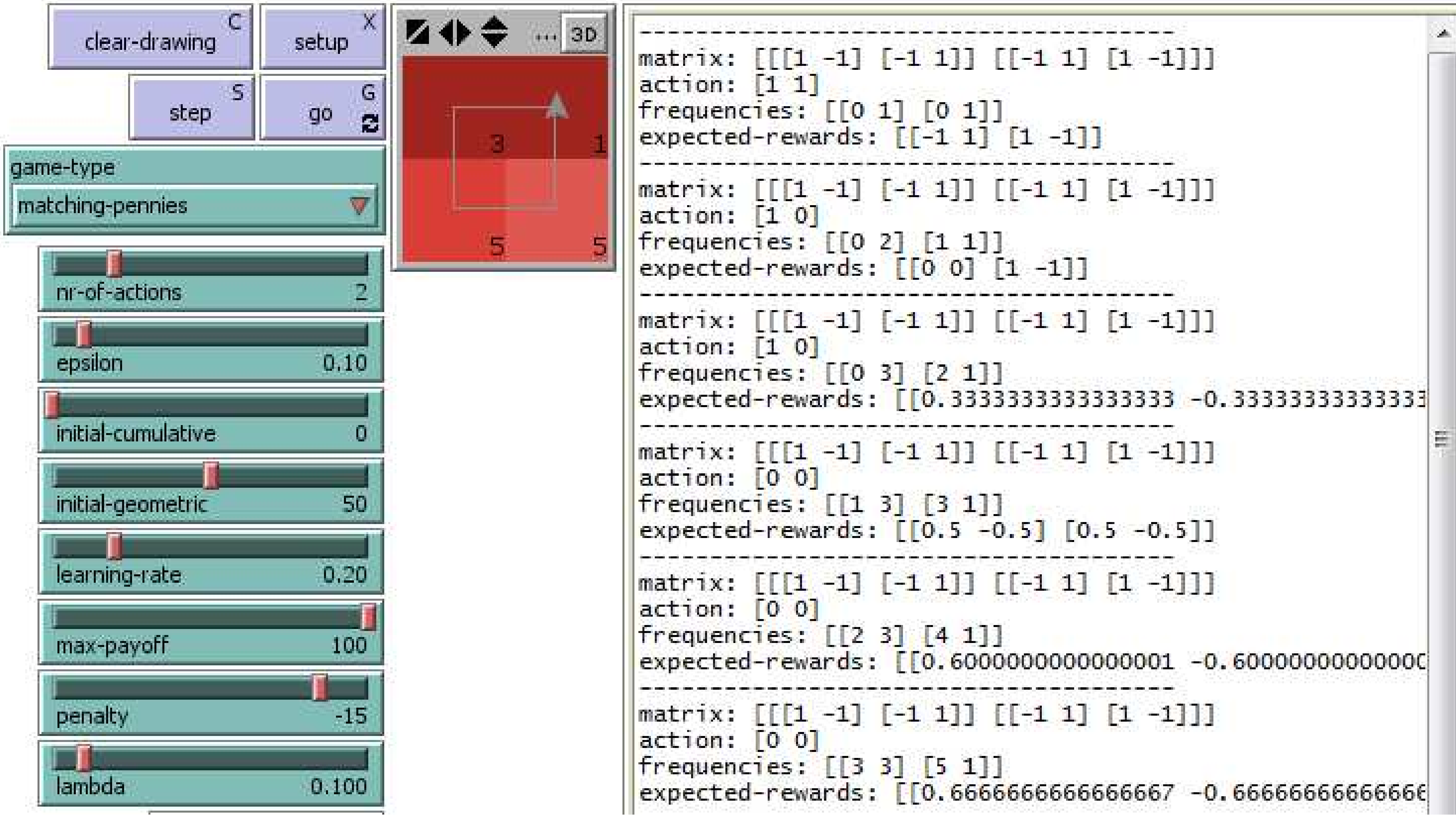
Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$
4.	H	H	$(4.5, 3.0)$	$(3.0, 4.5)$
5.	H	H	$(5.5, 3.0)$	$(4.0, 4.5)$
6.	H	H	$(6.5, 3.0)$	$(5.0, 4.5)$
7.	H	T	$(6.5, 4.0)$	$(6.0, 4.5)$
8.	H	T	$(6.5, 5.0)$	$(7.0, 4.5)$

What if all equilibria are mixed?

Example. Matching Pennies. This is a zero-sum game. A's goal is to have pennies matched. B's goal is the opposite.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(1.5, 2.0)$	$(2.0, 1.5)$
1.	T	T	$(1.5, 3.0)$	$(2.0, 2.5)$
2.	T	H	$(2.5, 3.0)$	$(2.0, 3.5)$
3.	T	H	$(3.5, 3.0)$	$(2.0, 4.5)$
4.	H	H	$(4.5, 3.0)$	$(3.0, 4.5)$
5.	H	H	$(5.5, 3.0)$	$(4.0, 4.5)$
6.	H	H	$(6.5, 3.0)$	$(5.0, 4.5)$
7.	H	T	$(6.5, 4.0)$	$(6.0, 4.5)$
8.	H	T	$(6.5, 5.0)$	$(7.0, 4.5)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Frequencies of fictitious play



Convergent empirical distribution of strategies

Theorem. *If the empirical distribution of strategies converges in fictitious play, then it converges to a Nash equilibrium.*

Convergent empirical distribution of strategies

Theorem. *If the empirical distribution of strategies converges in fictitious play, then it converges to a Nash equilibrium.*

Proof.

Convergent empirical distribution of strategies

Theorem. *If the empirical distribution of strategies converges in fictitious play, then it converges to a Nash equilibrium.*

Proof. Let i be arbitrary.

Convergent empirical distribution of strategies

Theorem. *If the empirical distribution of strategies converges in fictitious play, then it converges to a Nash equilibrium.*

Proof. Let i be arbitrary. If the empirical distribution converges to q, then i's opponent model converges to q^{-i}.

Convergent empirical distribution of strategies

Theorem. *If the empirical distribution of strategies converges in fictitious play, then it converges to a Nash equilibrium.*

Proof. Let i be arbitrary. If the empirical distribution converges to q, then i's opponent model converges to q^{-i}. By definition of fictitious play, $q^i \in \mathrm{BR}(q^{-i})$.

Convergent empirical distribution of strategies

Theorem. *If the empirical distribution of strategies converges in fictitious play, then it converges to a Nash equilibrium.*

Proof. Let i be arbitrary. If the empirical distribution converges to q, then i's opponent model converges to q^{-i}. By definition of fictitious play, $q^i \in \mathrm{BR}(q^{-i})$. Because of convergence, all such (mixed) best replies converge along.

Convergent empirical distribution of strategies

Theorem. *If the empirical distribution of strategies converges in fictitious play, then it converges to a Nash equilibrium.*

Proof. Let i be arbitrary. If the empirical distribution converges to q, then i's opponent model converges to q^{-i}. By definition of fictitious play, $q^i \in \mathrm{BR}(q^{-i})$. Because of convergence, all such (mixed) best replies converge along. By definition q is a Nash equilibrium. $\square$

Convergent empirical distribution of strategies

Theorem. *If the empirical distribution of strategies converges in fictitious play, then it converges to a Nash equilibrium.*

Proof. Let i be arbitrary. If the empirical distribution converges to q, then i's opponent model converges to q^{-i}. By definition of fictitious play, $q^i \in \mathrm{BR}(q^{-i})$. Because of convergence, all such (mixed) best replies converge along. By definition q is a Nash equilibrium. $\square$

Remarks:

Convergent empirical distribution of strategies

Theorem. *If the empirical distribution of strategies converges in fictitious play, then it converges to a Nash equilibrium.*

Proof. Let i be arbitrary. If the empirical distribution converges to q, then i's opponent model converges to q^{-i}. By definition of fictitious play, $q^i \in \mathrm{BR}(q^{-i})$. Because of convergence, all such (mixed) best replies converge along. By definition q is a Nash equilibrium. $\square$

Remarks:

1. The q^i may be mixed.

Convergent empirical distribution of strategies

Theorem. *If the empirical distribution of strategies converges in fictitious play, then it converges to a Nash equilibrium.*

Proof. Let i be arbitrary. If the empirical distribution converges to q, then i's opponent model converges to q^{-i}. By definition of fictitious play, $q^i \in \mathrm{BR}(q^{-i})$. Because of convergence, all such (mixed) best replies converge along. By definition q is a Nash equilibrium. $\square$

Remarks:

1. The q^i may be mixed.

2. It actually suffices that the q^{-i} converge asymptotically to the actual distribution (Fudenberg & Levine, 1998).

Convergent empirical distribution of strategies

Theorem. *If the empirical distribution of strategies converges in fictitious play, then it converges to a Nash equilibrium.*

Proof. Let i be arbitrary. If the empirical distribution converges to q, then i's opponent model converges to q^{-i}. By definition of fictitious play, $q^i \in \text{BR}(q^{-i})$. Because of convergence, all such (mixed) best replies converge along. By definition q is a Nash equilibrium. $\square$

Remarks:

1. The q^i may be mixed.
2. It actually suffices that the q^{-i} converge asymptotically to the actual distribution (Fudenberg & Levine, 1998).
3. If the empirical distributions converge (hence, converge to a Nash equilibrium), the actually played responses per stage need not be Nash equilibria of the stage game.

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	**A's action**	**B's action**	**A's beliefs**	**B's beliefs**

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A's action	B's action	A's beliefs	B's beliefs
0.				

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0.5, 1.0)$	$(1.0, 0.5)$

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0.5, 1.0)$	$(1.0, 0.5)$
1.	B	A		

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0.5, 1.0)$	$(1.0, 0.5)$
1.	B	A	$(1.5, 1.0)$	$(1.0, 1.5)$

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0.5, 1.0)$	$(1.0, 0.5)$
1.	B	A	$(1.5, 1.0)$	$(1.0, 1.5)$
2.	A	B		

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0.5, 1.0)$	$(1.0, 0.5)$
1.	B	A	$(1.5, 1.0)$	$(1.0, 1.5)$
2.	A	B	$(1.5, 2.0)$	$(2.0, 1.5)$

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0.5, 1.0)$	$(1.0, 0.5)$
1.	B	A	$(1.5, 1.0)$	$(1.0, 1.5)$
2.	A	B	$(1.5, 2.0)$	$(2.0, 1.5)$
3.	B	A		

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0.5, 1.0)$	$(1.0, 0.5)$
1.	B	A	$(1.5, 1.0)$	$(1.0, 1.5)$
2.	A	B	$(1.5, 2.0)$	$(2.0, 1.5)$
3.	B	A	$(2.5, 2.0)$	$(2.0, 2.5)$

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0.5, 1.0)$	$(1.0, 0.5)$
1.	B	A	$(1.5, 1.0)$	$(1.0, 1.5)$
2.	A	B	$(1.5, 2.0)$	$(2.0, 1.5)$
3.	B	A	$(2.5, 2.0)$	$(2.0, 2.5)$
4.	A	B		

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0.5, 1.0)$	$(1.0, 0.5)$
1.	B	A	$(1.5, 1.0)$	$(1.0, 1.5)$
2.	A	B	$(1.5, 2.0)$	$(2.0, 1.5)$
3.	B	A	$(2.5, 2.0)$	$(2.0, 2.5)$
4.	A	B	$(2.5, 3.0)$	$(3.0, 2.5)$

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0.5, 1.0)$	$(1.0, 0.5)$
1.	B	A	$(1.5, 1.0)$	$(1.0, 1.5)$
2.	A	B	$(1.5, 2.0)$	$(2.0, 1.5)$
3.	B	A	$(2.5, 2.0)$	$(2.0, 2.5)$
4.	A	B	$(2.5, 3.0)$	$(3.0, 2.5)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0.5, 1.0)$	$(1.0, 0.5)$
1.	B	A	$(1.5, 1.0)$	$(1.0, 1.5)$
2.	A	B	$(1.5, 2.0)$	$(2.0, 1.5)$
3.	B	A	$(2.5, 2.0)$	$(2.0, 2.5)$
4.	A	B	$(2.5, 3.0)$	$(3.0, 2.5)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

- This game possesses three equilibria, viz. $(0, 0)$, $(0.5, 0.5)$, and $(1, 1)$, with expected payoffs 1, 0.5, and 1, respectively.

Empirical distributions converge to Nash $\not\Rightarrow$ stage Nash

Repeated Coordination Game. Players receive payoff 1 iff they coordinate, else 0.

Round	***A*'s action**	***B*'s action**	***A*'s beliefs**	***B*'s beliefs**
0.			$(0.5, 1.0)$	$(1.0, 0.5)$
1.	B	A	$(1.5, 1.0)$	$(1.0, 1.5)$
2.	A	B	$(1.5, 2.0)$	$(2.0, 1.5)$
3.	B	A	$(2.5, 2.0)$	$(2.0, 2.5)$
4.	A	B	$(2.5, 3.0)$	$(3.0, 2.5)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

- This game possesses three equilibria, viz. $(0, 0)$, $(0.5, 0.5)$, and $(1, 1)$, with expected payoffs 1, 0.5, and 1, respectively.
- Empirical distribution of play converges to $(0.5, 0.5)$,—with payoff 0, rather than 0.5.

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.				

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	***A*'s action**	***B*'s action**	***A*'s beliefs**	***B*'s beliefs**
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$
1.	Rock	Scissors		

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	**A's action**	**B's action**	**A's beliefs**	**B's beliefs**
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$
1.	Rock	Scissors	$(0.0, 0.0, 1.5)$	$(1.0, 0.5, 0.0)$

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$
1.	Rock	Scissors	$(0.0, 0.0, 1.5)$	$(1.0, 0.5, 0.0)$
2.	Rock	Paper		

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$
1.	Rock	Scissors	$(0.0, 0.0, 1.5)$	$(1.0, 0.5, 0.0)$
2.	Rock	Paper	$(0.0, 1.0, 1.5)$	$(2.0, 0.5, 0.0)$

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$
1.	Rock	Scissors	$(0.0, 0.0, 1.5)$	$(1.0, 0.5, 0.0)$
2.	Rock	Paper	$(0.0, 1.0, 1.5)$	$(2.0, 0.5, 0.0)$
3.	Rock	Paper		

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$
1.	Rock	Scissors	$(0.0, 0.0, 1.5)$	$(1.0, 0.5, 0.0)$
2.	Rock	Paper	$(0.0, 1.0, 1.5)$	$(2.0, 0.5, 0.0)$
3.	Rock	Paper	$(0.0, 2.0, 1.5)$	$(3.0, 0.5, 0.0)$

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$
1.	Rock	Scissors	$(0.0, 0.0, 1.5)$	$(1.0, 0.5, 0.0)$
2.	Rock	Paper	$(0.0, 1.0, 1.5)$	$(2.0, 0.5, 0.0)$
3.	Rock	Paper	$(0.0, 2.0, 1.5)$	$(3.0, 0.5, 0.0)$
4.	Scissors	Paper		

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$
1.	Rock	Scissors	$(0.0, 0.0, 1.5)$	$(1.0, 0.5, 0.0)$
2.	Rock	Paper	$(0.0, 1.0, 1.5)$	$(2.0, 0.5, 0.0)$
3.	Rock	Paper	$(0.0, 2.0, 1.5)$	$(3.0, 0.5, 0.0)$
4.	Scissors	Paper	$(0.0, 3.0, 1.5)$	$(3.0, 0.5, 1.0)$

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A's action	B's action	A's beliefs	B's beliefs
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$
1.	Rock	Scissors	$(0.0, 0.0, 1.5)$	$(1.0, 0.5, 0.0)$
2.	Rock	Paper	$(0.0, 1.0, 1.5)$	$(2.0, 0.5, 0.0)$
3.	Rock	Paper	$(0.0, 2.0, 1.5)$	$(3.0, 0.5, 0.0)$
4.	Scissors	Paper	$(0.0, 3.0, 1.5)$	$(3.0, 0.5, 1.0)$
5.	Scissors	Paper		

Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$
1.	Rock	Scissors	$(0.0, 0.0, 1.5)$	$(1.0, 0.5, 0.0)$
2.	Rock	Paper	$(0.0, 1.0, 1.5)$	$(2.0, 0.5, 0.0)$
3.	Rock	Paper	$(0.0, 2.0, 1.5)$	$(3.0, 0.5, 0.0)$
4.	Scissors	Paper	$(0.0, 3.0, 1.5)$	$(3.0, 0.5, 1.0)$
5.	Scissors	Paper	$(0.0, 4.0, 1.5)$	$(3.0, 0.5, 2.0)$

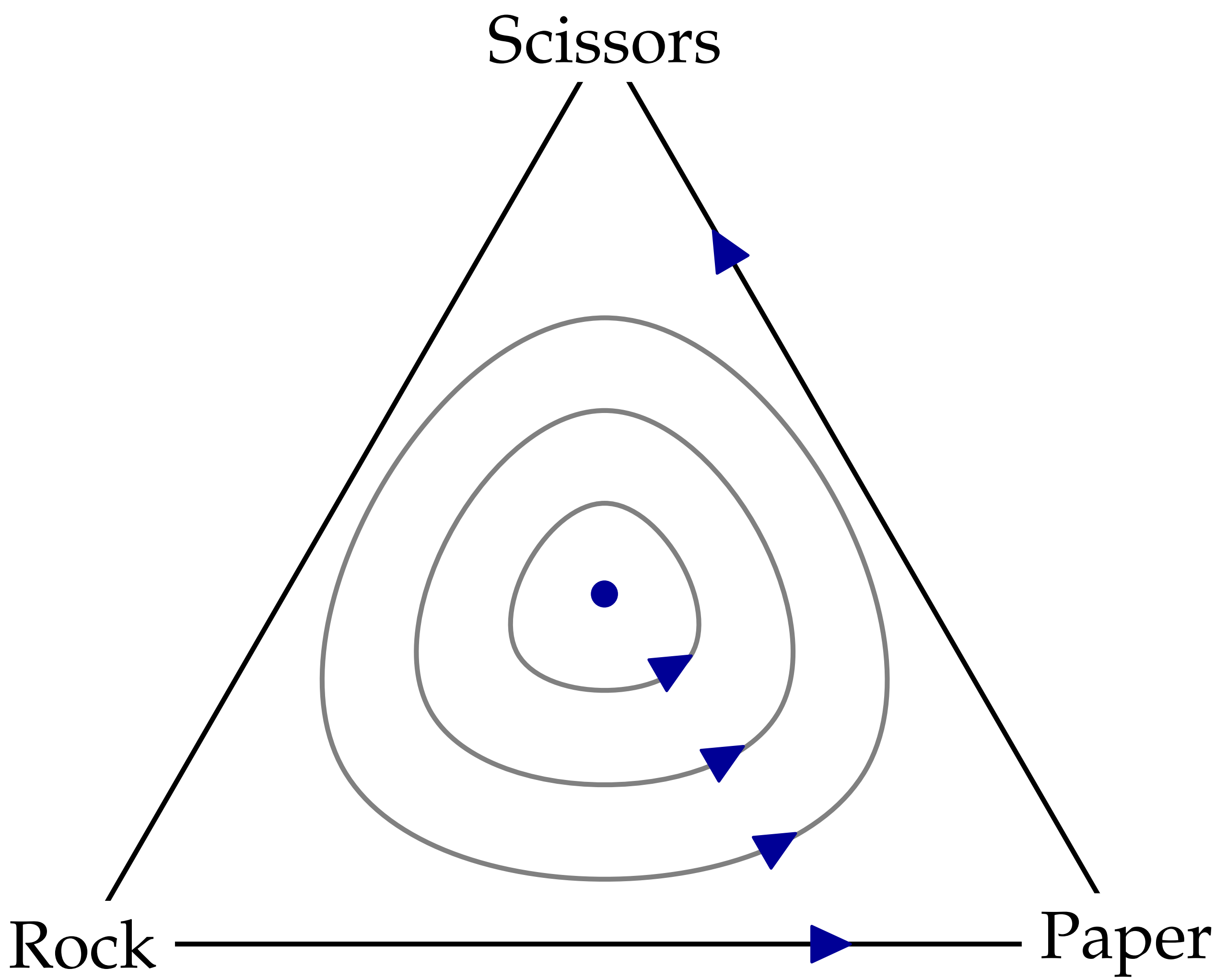
Empirical distr. of play does not need to converge

Rock-paper-scissors. Winner receives payoff 1, else 0.

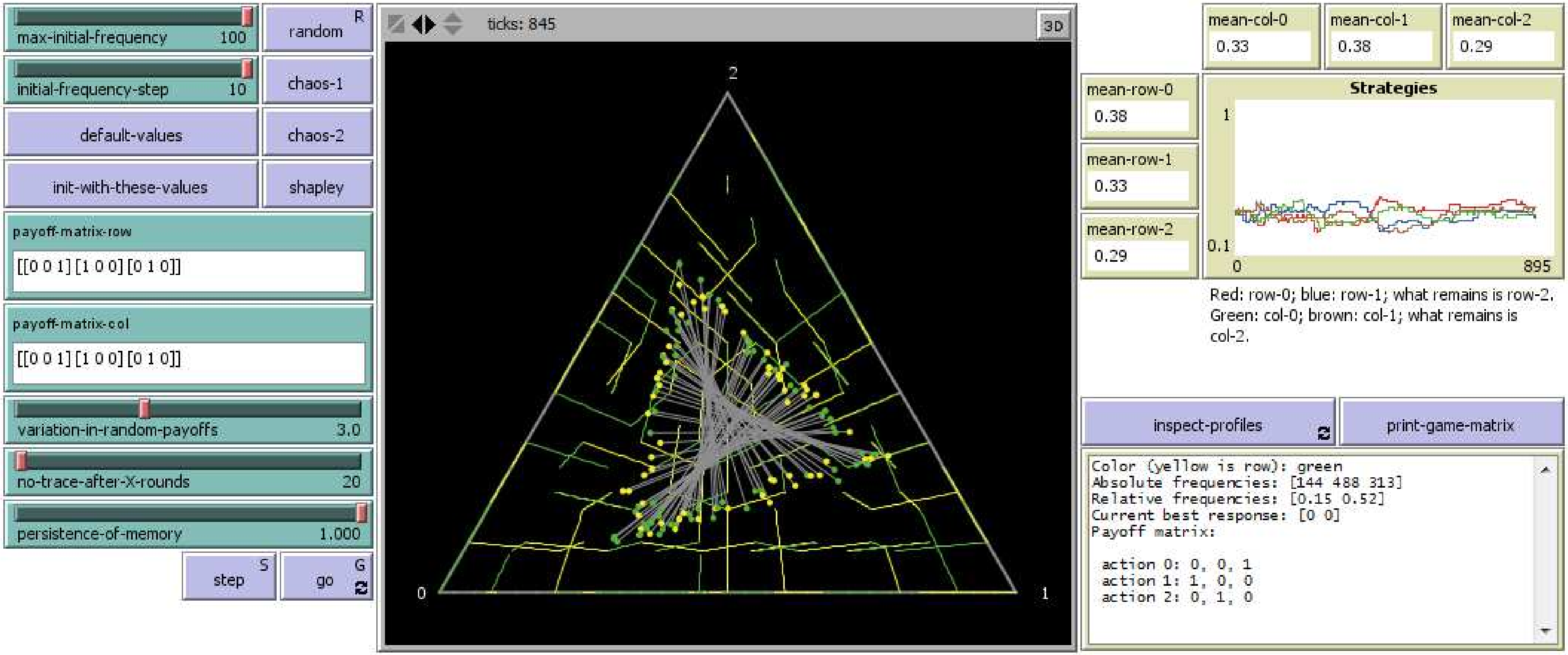
- Rock-paper-scissors with these payoffs is known as Shapley's game.
- Shapley's game possesses one equilibrium, viz. $(1/3, 1/3, 1/3)$, with expected payoff $1/3$.

Round	A**'s action**	B**'s action**	A**'s beliefs**	B**'s beliefs**
0.			$(0.0, 0.0, 0.5)$	$(0.0, 0.5, 0.0)$
1.	Rock	Scissors	$(0.0, 0.0, 1.5)$	$(1.0, 0.5, 0.0)$
2.	Rock	Paper	$(0.0, 1.0, 1.5)$	$(2.0, 0.5, 0.0)$
3.	Rock	Paper	$(0.0, 2.0, 1.5)$	$(3.0, 0.5, 0.0)$
4.	Scissors	Paper	$(0.0, 3.0, 1.5)$	$(3.0, 0.5, 1.0)$
5.	Scissors	Paper	$(0.0, 4.0, 1.5)$	$(3.0, 0.5, 2.0)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Repeated Shapley game: phase diagram

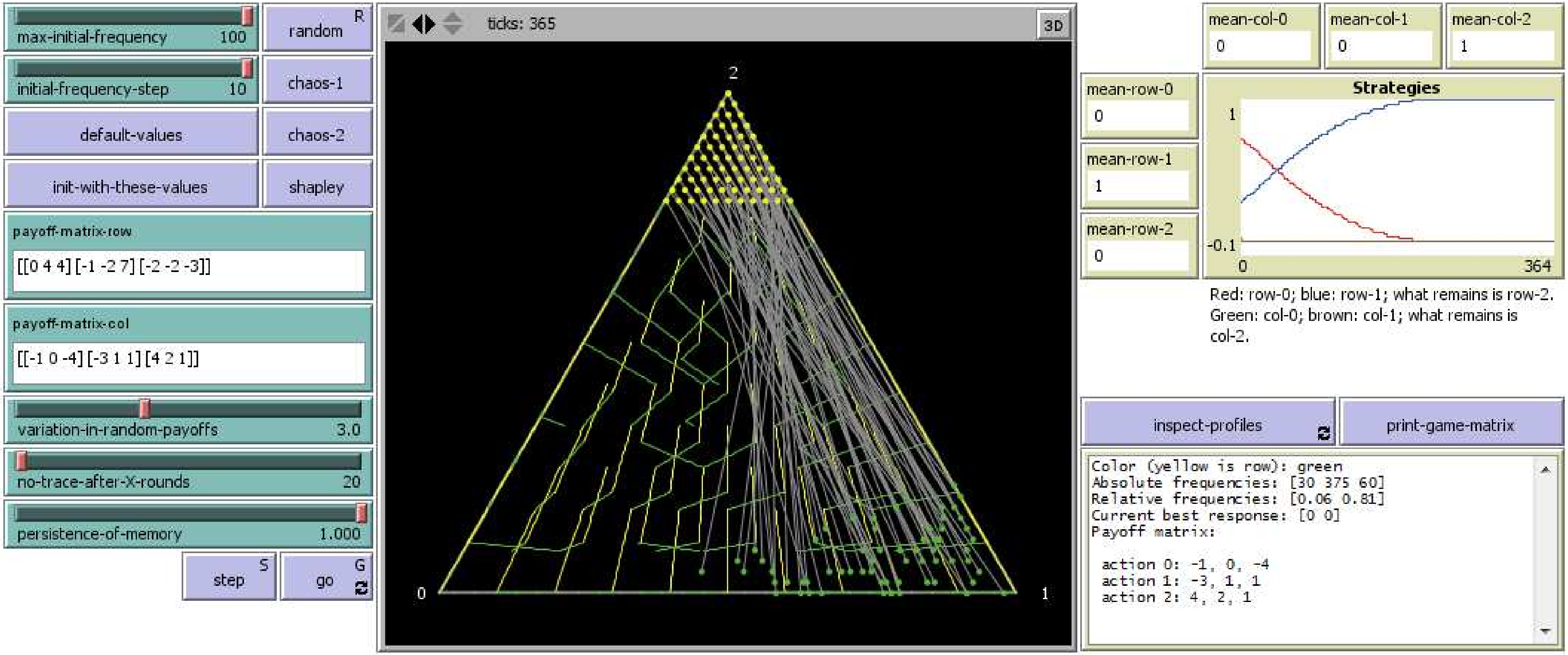


FP on Shapley’s game; strategy profiles in a simplex



There are many player couples. Each couple is connected by a gray line. Yellow is row; green is column. Player location is determined by the mixed strategy it projects on its opponent (i.e., normalised action count of its opponent). Each player starts with a biased action count. For example, with $[100, 0, 0]$ (lower left) or $[0, 100, 0]$ (lower right) or $[33, 33, 33]$ (center). Initial action counts of player pairs are unrelated.

FP on a 3x3 game; strategy profiles in a simplex



There are many player couples. Each couple is connected by a gray line. Yellow is row; green is column. Player location is determined by the mixed strategy it projects on its opponent (i.e., normalised action count of its opponent). Each player starts with a biased action count. For example, with $[100, 0, 0]$ (lower left) or $[0, 100, 0]$ (lower right) or $[33, 33, 33]$ (center). Initial action counts of player pairs are unrelated.

Part II:
Extensions and approximations of fictitious play

Proposed extensions to fictitious play

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**, say on m most recent rounds.

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**, say on m most recent rounds.
   - **Discounted data**

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**, say on m most recent rounds.
   - **Discounted data**, say with discount factor γ.

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**, say on m most recent rounds.
   - **Discounted data**, say with discount factor γ.
   - **Perturbed data**

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**, say on m most recent rounds.
   - **Discounted data**, say with discount factor γ.
   - **Perturbed data**, say with error ϵ on individual observations.

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**, say on m most recent rounds.
   - **Discounted data**, say with discount factor γ.
   - **Perturbed data**, say with error ϵ on individual observations.
   - **Random samples** of historical data

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**, say on m most recent rounds.
   - **Discounted data**, say with discount factor γ.
   - **Perturbed data**, say with error ϵ on individual observations.
   - **Random samples** of historical data, say on random samples of size m.

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**, say on m most recent rounds.
   - **Discounted data**, say with discount factor γ.
   - **Perturbed data**, say with error ϵ on individual observations.
   - **Random samples** of historical data, say on random samples of size m.
2. Give not necessarily *best responses*, but

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**, say on m most recent rounds.
   - **Discounted data**, say with discount factor γ.
   - **Perturbed data**, say with error ϵ on individual observations.
   - **Random samples** of historical data, say on random samples of size m.
2. Give not necessarily *best responses*, but
   - ϵ-**greedy**.

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**, say on m most recent rounds.
   - **Discounted data**, say with discount factor γ.
   - **Perturbed data**, say with error ϵ on individual observations.
   - **Random samples** of historical data, say on random samples of size m.
2. Give not necessarily *best responses*, but
   - ϵ**-greedy**.
   - **Perturbed throughout**, with small random shocks.

Proposed extensions to fictitious play

1. Build forecasts, not on *complete history*, but on
   - **Recent data**, say on m most recent rounds.
   - **Discounted data**, say with discount factor γ.
   - **Perturbed data**, say with error ϵ on individual observations.
   - **Random samples** of historical data, say on random samples of size m.
2. Give not necessarily *best responses*, but
   - ϵ**-greedy**.
   - **Perturbed throughout**, with small random shocks.
   - Randomly, and **proportional to expected payoff**.

Jordan’s framework for FP

Framework for predictive learning (e.g. fictitious play)

Framework for predictive learning (e.g. fictitious play)

A forecasting rule for player i is a function that maps a history to a probability distribution over the opponents' actions in the next round

Framework for predictive learning (e.g. fictitious play)

A forecasting rule for player i is a function that maps a history to a probability distribution over the opponents' actions in the next round:

$$f_i : H \to \Delta(X_{-i}).$$

Framework for predictive learning (e.g. fictitious play)

A forecasting rule for player i is a function that maps a history to a probability distribution over the opponents' actions in the next round:

$$f_i : H \rightarrow \Delta(X_{-i}).$$

A response rule for player i is a function that maps a history to a probability distribution over i's own actions in the next round

Framework for predictive learning (e.g. fictitious play)

A forecasting rule for player i is a function that maps a history to a probability distribution over the opponents' actions in the next round:

$$f_i : H \to \Delta(X_{-i}).$$

A response rule for player i is a function that maps a history to a probability distribution over i's own actions in the next round :

$$g_i : H \to \Delta(X_i).$$

Framework for predictive learning (e.g. fictitious play)

A forecasting rule for player i is a function that maps a history to a probability distribution over the opponents' actions in the next round:

$$f_i : H \to \Delta(X_{-i}).$$

A response rule for player i is a function that maps a history to a probability distribution over i's own actions in the next round :

$$g_i : H \to \Delta(X_i).$$

A predictive learning rule for player i is the combination of a forecasting rule and a response rule.

Framework for predictive learning (e.g. fictitious play)

A forecasting rule for player i is a function that maps a history to a probability distribution over the opponents' actions in the next round:

$$f_i : H \to \Delta(X_{-i}).$$

A response rule for player i is a function that maps a history to a probability distribution over i's own actions in the next round :

$$g_i : H \to \Delta(X_i).$$

A predictive learning rule for player i is the combination of a forecasting rule and a response rule. This is typically written as (f_i, g_i).

Framework for predictive learning (e.g. fictitious play)

A forecasting rule for player i is a function that maps a history to a probability distribution over the opponents' actions in the next round:

$$f_i : H \to \Delta(X_{-i}).$$

A response rule for player i is a function that maps a history to a probability distribution over i's own actions in the next round :

$$g_i : H \to \Delta(X_i).$$

A predictive learning rule for player i is the combination of a forecasting rule and a response rule. This is typically written as (f_i, g_i).

- This framework can be attributed to J.S. Jordan (1993).

Framework for predictive learning (e.g. fictitious play)

A forecasting rule for player i is a function that maps a history to a probability distribution over the opponents' actions in the next round:

$$f_i : H \to \Delta(X_{-i}).$$

A response rule for player i is a function that maps a history to a probability distribution over i's own actions in the next round :

$$g_i : H \to \Delta(X_i).$$

A predictive learning rule for player i is the combination of a forecasting rule and a response rule. This is typically written as (f_i, g_i).

- This framework can be attributed to J.S. Jordan (1993).
- Forecasting and response functions are deterministic.

Framework for predictive learning (e.g. fictitious play)

A forecasting rule for player i is a function that maps a history to a probability distribution over the opponents' actions in the next round:

$$f_i : H \rightarrow \Delta(X_{-i}).$$

A response rule for player i is a function that maps a history to a probability distribution over i's own actions in the next round :

$$g_i : H \rightarrow \Delta(X_i).$$

A predictive learning rule for player i is the combination of a forecasting rule and a response rule. This is typically written as (f_i, g_i).

- This framework can be attributed to J.S. Jordan (1993).
- Forecasting and response functions are deterministic.
- Reinforcement and regret do not fit.

Framework for predictive learning (e.g. fictitious play)

A forecasting rule for player i is a function that maps a history to a probability distribution over the opponents' actions in the next round:

$$f_i : H \to \Delta(X_{-i}).$$

A response rule for player i is a function that maps a history to a probability distribution over i's own actions in the next round :

$$g_i : H \to \Delta(X_i).$$

A predictive learning rule for player i is the combination of a forecasting rule and a response rule. This is typically written as (f_i, g_i).

- This framework can be attributed to J.S. Jordan (1993).
- Forecasting and response functions are deterministic.
- Reinforcement and regret do not fit. (They are not involved with prediction.)

Forecasting and response rules for fictitious play

Forecasting and response rules for fictitious play

Let $h^t \in H^t$ be a history of play *up to and including* round t

Forecasting and response rules for fictitious play

Let $h^t \in H^t$ be a history of play *up to and including* round t and

$$\phi^{jt} =_{Def} \text{the empirical distribution of } j\text{'s actions u.t.a.i. round } t.$$

Forecasting and response rules for fictitious play

Let $h^t \in H^t$ be a history of play *up to and including* round t and

$\phi^{jt} =_{Def}$ the empirical distribution of j's actions u.t.a.i. round t.

Then the fictitious play forecasting rule is given by

$$f_i(h^t) =_{Def} \prod_{j \neq i} \phi^{jt}.$$

Forecasting and response rules for fictitious play

Let $h^t \in H^t$ be a history of play *up to and including* round t and

$$\phi^{jt} =_{Def} \text{the empirical distribution of } j\text{'s actions u.t.a.i. round } t.$$

Then the fictitious play forecasting rule is given by

$$f_i(h^t) =_{Def} \prod_{j \neq i} \phi^{jt}.$$

Let f_i be a fictitious play forecasting rule.

Forecasting and response rules for fictitious play

Let $h^t \in H^t$ be a history of play *up to and including* round t and

$\phi^{jt} =_{Def}$ the empirical distribution of j's actions u.t.a.i. round t.

Then the fictitious play forecasting rule is given by

$$f_i(h^t) =_{Def} \prod_{j \neq i} \phi^{jt}.$$

Let f_i be a fictitious play forecasting rule. Then g_i is said to be a fictitious play response rule if all values are best responses to values of f_i.

Forecasting and response rules for fictitious play

Let $h^t \in H^t$ be a history of play *up to and including* round t and

$$\phi^{jt} =_{Def} \text{the empirical distribution of } j\text{'s actions u.t.a.i. round } t.$$

Then the fictitious play forecasting rule is given by

$$f_i(h^t) =_{Def} \prod_{j \neq i} \phi^{jt}.$$

Let f_i be a fictitious play forecasting rule. Then g_i is said to be a fictitious play response rule if all values are best responses to values of f_i.

Remarks:

Forecasting and response rules for fictitious play

Let $h^t \in H^t$ be a history of play *up to and including* round t and

$\phi^{jt} =_{Def}$ the empirical distribution of j's actions u.t.a.i. round t.

Then the fictitious play forecasting rule is given by

$$f_i(h^t) =_{Def} \prod_{j \neq i} \phi^{jt}.$$

Let f_i be a fictitious play forecasting rule. Then g_i is said to be a fictitious play response rule if all values are best responses to values of f_i.

Remarks:

1. Player i attributes a mixed strategy ϕ^{jt} to player j. This strategy reflects the number of times each action is played by j.

Forecasting and response rules for fictitious play

Let $h^t \in H^t$ be a history of play *up to and including* round t and

$$\phi^{jt} =_{Def} \text{the empirical distribution of } j\text{'s actions u.t.a.i. round } t.$$

Then the fictitious play forecasting rule is given by

$$f_i(h^t) =_{Def} \prod_{j \neq i} \phi^{jt}.$$

Let f_i be a fictitious play forecasting rule. Then g_i is said to be a fictitious play response rule if all values are best responses to values of f_i.

Remarks:

1. Player i attributes a mixed strategy ϕ^{jt} to player j. This strategy reflects the number of times each action is played by j.
2. The mixed strategies are assumed to be independent.

Forecasting and response rules for fictitious play

Let $h^t \in H^t$ be a history of play *up to and including* round t and

$$\phi^{jt} =_{Def} \text{the empirical distribution of } j\text{'s actions u.t.a.i. round } t.$$

Then the fictitious play forecasting rule is given by

$$f_i(h^t) =_{Def} \prod_{j \neq i} \phi^{jt}.$$

Let f_i be a fictitious play forecasting rule. Then g_i is said to be a fictitious play response rule if all values are best responses to values of f_i.

Remarks:

1. Player i attributes a mixed strategy ϕ^{jt} to player j. This strategy reflects the number of times each action is played by j.
2. The mixed strategies are assumed to be independent.
3. Both (1) and (2) are simplifying assumptions.

Smoothed fictitious play

Smoothed fictitious play

Smoothed fictitious play

Notation:

Smoothed fictitious play

Notation:

p^{-i} : strategy profile of opponents as predicted by f_i in round t.

Smoothed fictitious play

Notation:

p^{-i} : strategy profile of opponents as predicted by f_i in round t.

$u_i(x_i, p^{-i})$: expected utility of action x_i, given p^{-i}.

Smoothed fictitious play

Notation:

p^{-i} : strategy profile of opponents as predicted by f_i in round t.

$u_i(x_i, p^{-i})$: expected utility of action x_i, given p^{-i}.

q^i : strategy profile of player i in round $t + 1$. I.e., $g_i(h)$.

Smoothed fictitious play

Notation:

p^{-i} : strategy profile of opponents as predicted by f_i in round t.
$u_i(x_i, p^{-i})$: expected utility of action x_i, given p^{-i}.
q^i : strategy profile of player i in round $t+1$. I.e., $g_i(h)$.

Task: define q^i given p^{-i} and $u_i(x_i, p^{-i})$.

Smoothed fictitious play

Notation:

p^{-i} : strategy profile of opponents as predicted by f_i in round t.
$u_i(x_i, p^{-i})$: expected utility of action x_i, given p^{-i}.
q^i : strategy profile of player i in round $t+1$. I.e., $g_i(h)$.

Task: define q^i given p^{-i} and $u_i(x_i, p^{-i})$.

Idea:

Respond randomly, but (somehow) proportional to expected payoff.

Smoothed fictitious play

Notation:

p^{-i} : strategy profile of opponents as predicted by f_i in round t.

$u_i(x_i, p^{-i})$: expected utility of action x_i, given p^{-i}.

q^i : strategy profile of player i in round $t+1$. I.e., $g_i(h)$.

Task: define q^i given p^{-i} and $u_i(x_i, p^{-i})$.

Idea:

Respond randomly, but (somehow) proportional to expected payoff.

Elaborations of this idea:

b) Through soft max (a.k.a. mixed logit):

$$q^i(x_i \mid p^{-i}) =_{Def} \frac{e^{u_i(x_i, p^{-i})/\gamma}}{\sum_{x'_i \in X_i} e^{u_i(x'_i, p^{-i})/\gamma}}.$$

Smoothed fictitious play

Notation:

p^{-i} : strategy profile of opponents as predicted by f_i in round t.
$u_i(x_i, p^{-i})$: expected utility of action x_i, given p^{-i}.
q^i : strategy profile of player i in round $t+1$. I.e., $g_i(h)$.

Task: define q^i given p^{-i} and $u_i(x_i, p^{-i})$.

Idea:

Respond randomly, but (somehow) proportional to expected payoff.

Elaborations of this idea:
a) Strictly proportional:

$$q^i(x_i \mid p^{-i}) =_{Def} \frac{u_i(x_i, p^{-i})}{\sum_{x'_i \in X_i} u_i(x'_i, p^{-i})}.$$

Smoothed fictitious play

Notation:

p^{-i} : strategy profile of opponents as predicted by f_i in round t.

$u_i(x_i, p^{-i})$: expected utility of action x_i, given p^{-i}.

q^i : strategy profile of player i in round $t+1$. I.e., $g_i(h)$.

Task: define q^i given p^{-i} and $u_i(x_i, p^{-i})$.

Idea:

> Respond randomly, but (somehow) proportional to expected payoff.

Elaborations of this idea:

a) Strictly proportional:

$$q^i(x_i \mid p^{-i}) =_{Def} \frac{u_i(x_i, p^{-i})}{\sum_{x'_i \in X_i} u_i(x'_i, p^{-i})}.$$

b) Through soft max (a.k.a. mixed logit):

$$q^i(x_i \mid p^{-i}) =_{Def} \frac{e^{u_i(x_i, p^{-i})/\gamma}}{\sum_{x'_i \in X_i} e^{u_i(x'_i, p^{-i})/\gamma}}.$$

Soft max a.k.a. mixed logit a.k.a. quantal response

Soft max:

$$q^i(x_i \mid p^{-i}) =_{Def} \frac{e^{u_i(x_i,p^{-i})/\gamma}}{\sum_{x'_i \in X_i} e^{u_i(x'_i,p^{-i})/\gamma}}.$$

Soft max a.k.a. mixed logit a.k.a. quantal response

Soft max:

$$q^i(x_i \mid p^{-i}) =_{Def} \frac{e^{u_i(x_i, p^{-i})/\gamma}}{\sum_{x_i' \in X_i} e^{u_i(x_i', p^{-i})/\gamma}}.$$

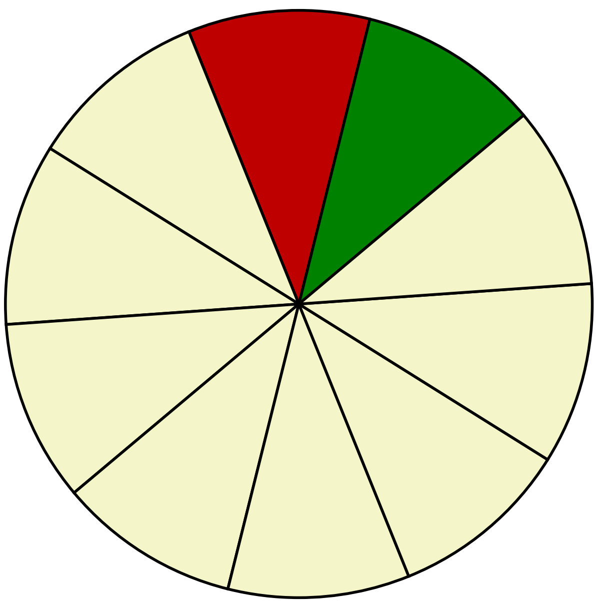
Soft max respects best replies, but leaves room for experimentation:

Soft max a.k.a. mixed logit a.k.a. quantal response

Soft max:

$$q^i(x_i \mid p^{-i}) =_{Def} \frac{e^{u_i(x_i,p^{-i})/\gamma}}{\sum_{x'_i \in X_i} e^{u_i(x'_i,p^{-i})/\gamma}}.$$

Soft max respects best replies, but leaves room for experimentation:



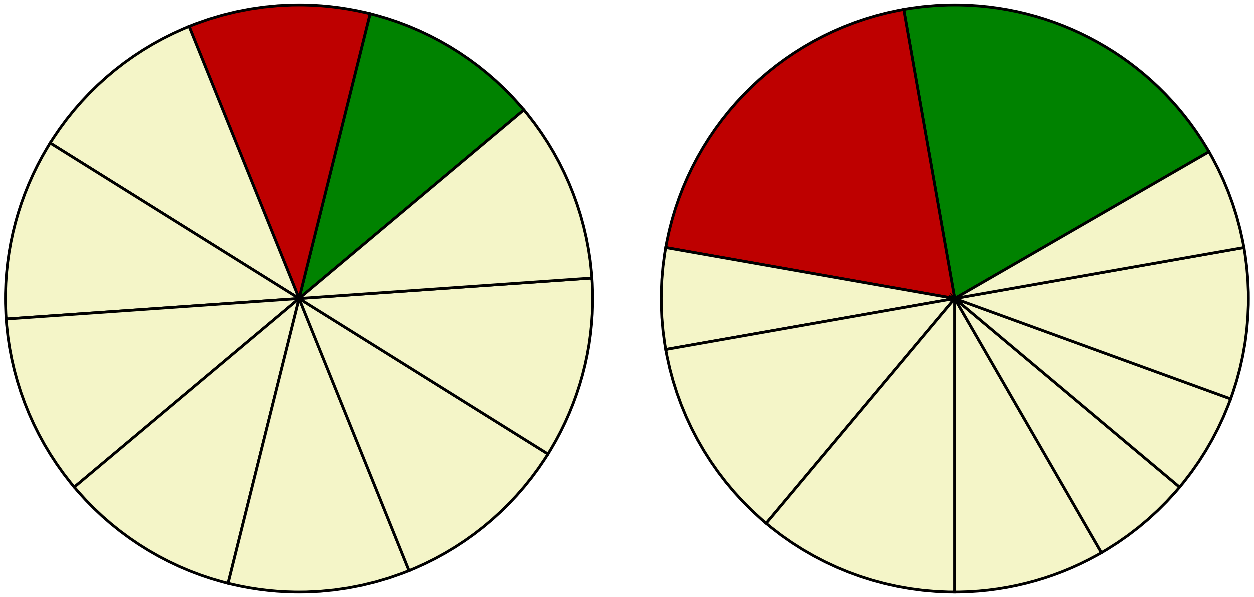
$\gamma \to \infty$

Soft max a.k.a. mixed logit a.k.a. quantal response

Soft max:

$$q^i(x_i \mid p^{-i}) =_{Def} \frac{e^{u_i(x_i, p^{-i})/\gamma}}{\sum_{x_i' \in X_i} e^{u_i(x_i', p^{-i})/\gamma}}.$$

Soft max respects best replies, but leaves room for experimentation:



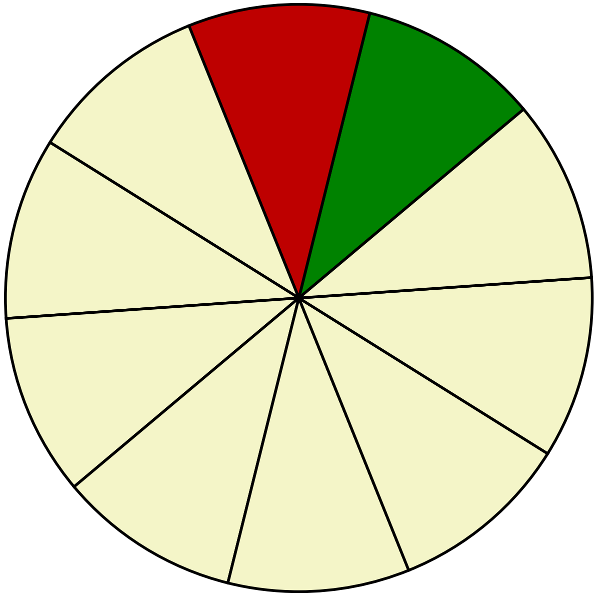
$\gamma \to \infty$ $\qquad$ $\gamma = 1$

Soft max a.k.a. mixed logit a.k.a. quantal response

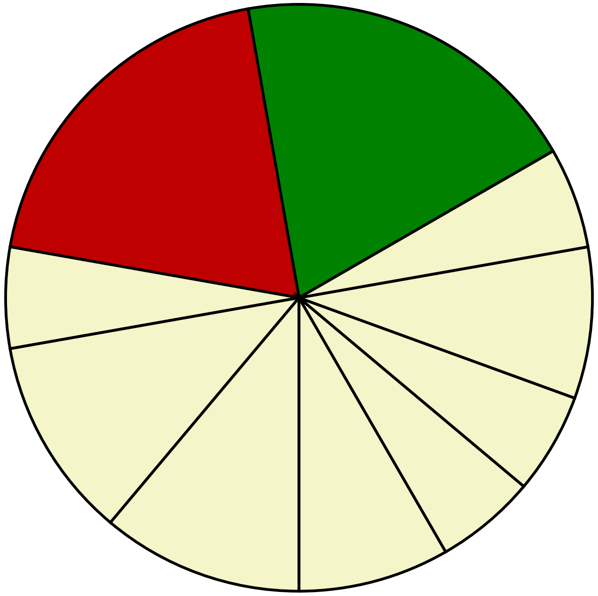
Soft max:

$$q^i(x_i \mid p^{-i}) =_{Def} \frac{e^{u_i(x_i, p^{-i})/\gamma}}{\sum_{x_i' \in X_i} e^{u_i(x_i', p^{-i})/\gamma}}.$$

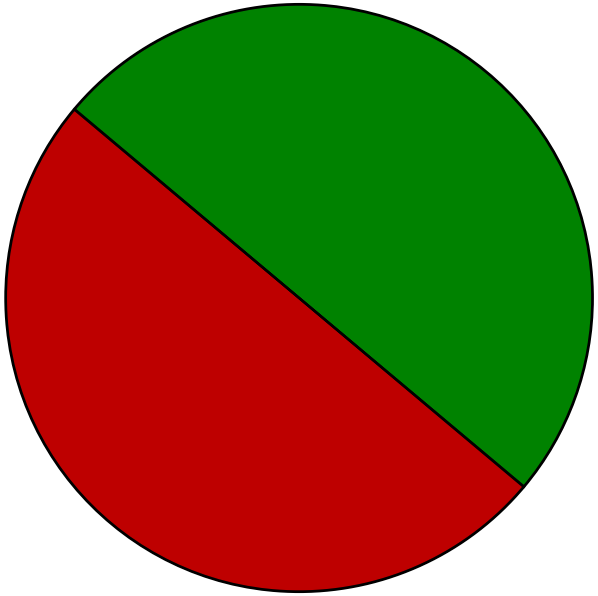
Soft max respects best replies, but leaves room for experimentation:



$\gamma \to \infty$



$\gamma = 1$



$\gamma \downarrow 0$

Smoothed fictitious play limits regret

Theorem (Fudenberg & Levine, 1995). *Let G be a finite game and let $\epsilon > 0$. If a given player uses smoothed fictitious play with a sufficiently small smoothing γ, then with probability one his regrets are bounded above by ϵ.*

Smoothed fictitious play limits regret

Theorem (Fudenberg & Levine, 1995). *Let G be a finite game and let $\epsilon > 0$. If a given player uses smoothed fictitious play with a sufficiently small smoothing γ, then with probability one his regrets are bounded above by ϵ.*

- Young does not reproduce the proof of Fudenberg *et al.*, but shows that in this case ϵ-regret can be derived from a later and more general result of Hart and Mas-Colell in 2001.

Smoothed fictitious play limits regret

Theorem (Fudenberg & Levine, 1995). *Let G be a finite game and let $\epsilon > 0$. If a given player uses smoothed fictitious play with a sufficiently small smoothing γ, then with probability one his regrets are bounded above by ϵ.*

- Young does not reproduce the proof of Fudenberg *et al.*, but shows that in this case ϵ-regret can be derived from a later and more general result of Hart and Mas-Colell in 2001.

- This later result identifies a large family of rules that eliminate regret, based on an extension of Blackwell's approachability theorem.

Smoothed fictitious play limits regret

Theorem (Fudenberg & Levine, 1995). *Let G be a finite game and let $\epsilon > 0$. If a given player uses smoothed fictitious play with a sufficiently small smoothing γ, then with probability one his regrets are bounded above by ϵ.*

- Young does not reproduce the proof of Fudenberg *et al.*, but shows that in this case ϵ-regret can be derived from a later and more general result of Hart and Mas-Colell in 2001.

- This later result identifies a large family of rules that eliminate regret, based on an extension of Blackwell's approachability theorem.

  (Blackwell's approachability theorem generalises maxmin reasoning to vector-valued payoffs.)

Smoothed fictitious play limits regret

Theorem (Fudenberg & Levine, 1995). *Let G be a finite game and let $\epsilon > 0$. If a given player uses smoothed fictitious play with a sufficiently small smoothing γ, then with probability one his regrets are bounded above by ϵ.*

- Young does not reproduce the proof of Fudenberg *et al.*, but shows that in this case ϵ-regret can be derived from a later and more general result of Hart and Mas-Colell in 2001.

- This later result identifies a large family of rules that eliminate regret, based on an extension of Blackwell's approachability theorem.

  (Blackwell's approachability theorem generalises maxmin reasoning to vector-valued payoffs.)

Fudenberg & Levine, 1995. "Consistency and cautious fictitious play," *Journal of Economic Dynamics and Control*, Vol. **19** (5-7), pp. 1065-1089.

Hart & Mas-Colell, 2001. "A General Class of Adaptive Strategies," *Journal of Economic Theory*, Vol. **98**(1), pp. 26-54.

Smoothed fictitious play converges to ϵ-CCE

Definition. Let X be action profiles, and $q \in \Delta(X)$. Then q is a coarse correlated equilibrium (CCE) if no one wants to opt out prior to a realisation of q in the form of an action profile.

In a coarse correlated ϵ-equilibrium (ϵ-CCE), no player wants to opt out to gain more in expected utility than ϵ.

Smoothed fictitious play converges to ϵ-CCE

Definition. Let X be action profiles, and $q \in \Delta(X)$. Then q is a coarse correlated equilibrium (CCE) if no one wants to opt out prior to a realisation of q in the form of an action profile.

In a coarse correlated ϵ-equilibrium (ϵ-CCE), no player wants to opt out to gain more in expected utility than ϵ.

Theorem (Fudenberg & Levine, 1995). *Let G be a finite game and let $\epsilon > 0$.*

Smoothed fictitious play converges to ϵ-CCE

Definition. Let X be action profiles, and $q \in \Delta(X)$. Then q is a coarse correlated equilibrium (CCE) if no one wants to opt out prior to a realisation of q in the form of an action profile.

In a coarse correlated ϵ-equilibrium (ϵ-CCE), no player wants to opt out to gain more in expected utility than ϵ.

Theorem (Fudenberg & Levine, 1995). *Let G be a finite game and let $\epsilon > 0$. If all players use smoothed fictitious play with sufficiently small smoothing parameters γ_i, then empirical play converges to the set of coarse correlated ϵ-equilibria a.s.*

Smoothed fictitious play converges to ϵ-CCE

Definition. Let X be action profiles, and $q \in \Delta(X)$. Then q is a coarse correlated equilibrium (CCE) if no one wants to opt out prior to a realisation of q in the form of an action profile.

In a coarse correlated ϵ-equilibrium (ϵ-CCE), no player wants to opt out to gain more in expected utility than ϵ.

Theorem (Fudenberg & Levine, 1995). *Let G be a finite game and let $\epsilon > 0$. If all players use smoothed fictitious play with sufficiently small smoothing parameters γ_i, then empirical play converges to the set of coarse correlated ϵ-equilibria a.s.*

Summary of the two theorems:

Smoothed fictitious play converges to ϵ-CCE

Definition. Let X be action profiles, and $q \in \Delta(X)$. Then q is a coarse correlated equilibrium (CCE) if no one wants to opt out prior to a realisation of q in the form of an action profile.

In a coarse correlated ϵ-equilibrium (ϵ-CCE), no player wants to opt out to gain more in expected utility than ϵ.

Theorem (Fudenberg & Levine, 1995). *Let G be a finite game and let $\epsilon > 0$. If all players use smoothed fictitious play with sufficiently small smoothing parameters γ_i, then empirical play converges to the set of coarse correlated ϵ-equilibria a.s.*

Summary of the two theorems:

- Smoothed fictitious play limits regret.

Smoothed fictitious play converges to ϵ-CCE

Definition. Let X be action profiles, and $q \in \Delta(X)$. Then q is a coarse correlated equilibrium (CCE) if no one wants to opt out prior to a realisation of q in the form of an action profile.

In a coarse correlated ϵ-equilibrium (ϵ-CCE), no player wants to opt out to gain more in expected utility than ϵ.

Theorem (Fudenberg & Levine, 1995). *Let G be a finite game and let $\epsilon > 0$. If all players use smoothed fictitious play with sufficiently small smoothing parameters γ_i, then empirical play converges to the set of coarse correlated ϵ-equilibria a.s.*

Summary of the two theorems:

- Smoothed fictitious play limits regret.
- Smoothed fictitious play converges to ϵ-CCE.

Smoothed fictitious play converges to ϵ-CCE

Definition. Let X be action profiles, and $q \in \Delta(X)$. Then q is a coarse correlated equilibrium (CCE) if no one wants to opt out prior to a realisation of q in the form of an action profile.

In a coarse correlated ϵ-equilibrium (ϵ-CCE), no player wants to opt out to gain more in expected utility than ϵ.

Theorem (Fudenberg & Levine, 1995). *Let G be a finite game and let $\epsilon > 0$. If all players use smoothed fictitious play with sufficiently small smoothing parameters γ_i, then empirical play converges to the set of coarse correlated ϵ-equilibria a.s.*

Summary of the two theorems:

- Smoothed fictitious play limits regret.
- Smoothed fictitious play converges to ϵ-CCE.

But there is another l.a. with **no** regret and convergence to **zero**-CCE!

Exponentiated regret matching

Exponentiated regret matching

Exponentiated regret matching

Let

Exponentiated regret matching

Let $\quad j$: action j, where $1 \leq j \leq k$

Exponentiated regret matching

Let j : action j, where $1 \leq j \leq k$

$\bar{u}_i^t$: i's (realised) average payoff up to and including round t

Exponentiated regret matching

Let

j : action j, where $1 \leq j \leq k$

$\bar{u}_i^t$: i's (realised) average payoff up to and including round t

ϕ^{-it} : the (realised) joint empirical distribution of i's opponents

Exponentiated regret matching

Let

$$
\begin{aligned}
j &: \text{action } j, \text{ where } 1 \leq j \leq k \\
\bar{u}_i^t &: i\text{'s (realised) average payoff up to and including round } t \\
\phi^{-it} &: \text{the (realised) joint empirical distribution of } i\text{'s opponents} \\
\bar{u}_i(j, \phi^{-it}) &: i\text{'s hypothetical mean payoff for playing } j \text{ against}
\end{aligned}
$$

Exponentiated regret matching

Let

$$
\begin{aligned}
j &: \text{action } j\text{, where } 1 \leq j \leq k \\
\bar{u}_i^t &: i\text{'s (realised) average payoff up to and including round } t \\
\phi^{-it} &: \text{the (realised) joint empirical distribution of } i\text{'s opponents} \\
\bar{u}_i(j, \phi^{-it}) &: i\text{'s hypothetical mean payoff for playing } j \text{ against } \phi^{-it}
\end{aligned}
$$

Exponentiated regret matching

Let

$$
\begin{aligned}
j &: \text{action } j\text{, where } 1 \leq j \leq k \\
\bar{u}_i^t &: i\text{'s (realised) average payoff up to and including round } t \\
\phi^{-it} &: \text{the (realised) joint empirical distribution of } i\text{'s opponents} \\
\bar{u}_i(j, \phi^{-it}) &: i\text{'s hypothetical mean payoff for playing } j \text{ against } \phi^{-it} \\
\bar{r}^{it} &: \text{player } i\text{'s regret vector in round } t\text{, i.e., } \bar{u}_i(j, \phi^{-it}) - \bar{u}_i^t
\end{aligned}
$$

Exponentiated regret matching

Let

$$\begin{aligned} j &: \text{action } j\text{, where } 1 \leq j \leq k \\ \bar{u}_i^t &: i\text{'s (realised) average payoff up to and including round } t \\ \phi^{-it} &: \text{the (realised) joint empirical distribution of } i\text{'s opponents} \\ \bar{u}_i(j, \phi^{-it}) &: i\text{'s hypothetical mean payoff for playing } j \text{ against } \phi^{-it} \\ \bar{r}^{it} &: \text{player } i\text{'s regret vector in round } t\text{, i.e., } \bar{u}_i(j, \phi^{-it}) - \bar{u}_i^t \end{aligned}$$

Exponentiated regret matching (PY, p. 59) is defined as

$$q_j^{i(t+1)} \propto [\bar{r}_j^{it}]_+^a$$

where $a > 0$.

Exponentiated regret matching

Let
j : action j, where $1 \leq j \leq k$
$\bar{u}_i^t$: i's (realised) average payoff up to and including round t
ϕ^{-it} : the (realised) joint empirical distribution of i's opponents
$\bar{u}_i(j, \phi^{-it})$: i's hypothetical mean payoff for playing j against ϕ^{-it}
$\bar{r}^{it}$: player i's regret vector in round t, i.e., $\bar{u}_i(j, \phi^{-it}) - \bar{u}_i^t$

Exponentiated regret matching (PY, p. 59) is defined as

$$q_j^{i(t+1)} \propto [\bar{r}_j^{it}]_+^a$$

where $a > 0$. (For $a = 1$ we have ordinary regret matching.)

Exponentiated regret matching

Let

$$\begin{aligned} j &: \text{action } j, \text{ where } 1 \leq j \leq k \\ \bar{u}_i^t &: i\text{'s (realised) average payoff up to and including round } t \\ \phi^{-it} &: \text{the (realised) joint empirical distribution of } i\text{'s opponents} \\ \bar{u}_i(j, \phi^{-it}) &: i\text{'s hypothetical mean payoff for playing } j \text{ against } \phi^{-it} \\ \bar{r}^{it} &: \text{player } i\text{'s regret vector in round } t, \text{ i.e., } \bar{u}_i(j, \phi^{-it}) - \bar{u}_i^t \end{aligned}$$

Exponentiated regret matching (PY, p. 59) is defined as

$$q_j^{i(t+1)} \propto [\bar{r}_j^{it}]_+^a$$

where $a > 0$. (For $a = 1$ we have ordinary regret matching.)

A theorem on exponentiated regret matching (Mas-Colell *et al.*, 2001) ensures that individual players have no-regret with probability one, and the empirical distribution of play converges to the set of coarse correlated equilibria (PY, p. 37 for RM, p. 60 for ERM).

FP

FP vs. Smoothed FP

FP vs. Smoothed FP vs. Exponentiated regret matching

Best reply vs. better reply

Best reply vs. better reply

- **Fictitious play** is a best reply strategy: it plays a best reply to sampled past play.

Best reply vs. better reply

- **Fictitious play** is a **best** reply strategy: it plays a best reply to sampled past play.
- **Regret matching** may be called a **better** reply strategy: it plays a reply for which the expected payoff is better than past average payoff.

Best reply vs. better reply

- **Fictitious play** is a **best** reply strategy: it plays a best reply to sampled past play.
- **Regret matching** may be called a **better** reply strategy: it plays a reply for which the expected payoff is better than past average payoff.
- **Smoothed fictitious play** is a best reply strategy nor a better reply strategy.

Best reply vs. better reply

- **Fictitious play** is a **best** reply strategy: it plays a best reply to sampled past play.
- **Regret matching** may be called a **better** reply strategy: it plays a reply for which the expected payoff is better than past average payoff.
- **Smoothed fictitious play** is a best reply strategy nor a better reply strategy.
  - It is not a best reply strategy because it puts positive probabilities on all actions at every period.

Best reply vs. better reply

- **Fictitious play** is a **best** reply strategy: it plays a best reply to sampled past play.
- **Regret matching** may be called a **better** reply strategy: it plays a reply for which the expected payoff is better than past average payoff.
- **Smoothed fictitious play** is a best reply strategy nor a better reply strategy.
  - It is not a best reply strategy because it puts positive probabilities on all actions at every period.
  - It is not a better reply strategy because it is not concerned with past payoffs.

Best reply vs. better reply

- **Fictitious play** is a **best** reply strategy: it plays a best reply to sampled past play.
- **Regret matching** may be called a **better** reply strategy: it plays a reply for which the expected payoff is better than past average payoff.
- **Smoothed fictitious play** is a best reply strategy nor a better reply strategy.
  - It is not a best reply strategy because it puts positive probabilities on all actions at every period.
  - It is not a better reply strategy because it is not concerned with past payoffs.

Hart, S., and Mas-Colell, A. (2000). "A simple adaptive procedure leading to correlated equilibrium". *Econometrica*, **68**, pp. 1127-1150.

FP vs. Smoothed FP vs. Exponentiated RM

FP vs. Smoothed FP vs. Exponentiated RM

Fictitious play Plays *best responses*.

FP vs. Smoothed FP vs. Exponentiated RM

Fictitious play Plays *best responses*.

- Does depend on past play of opponent(s).

FP vs. Smoothed FP vs. Exponentiated RM

Fictitious play Plays *best responses*.

- Does depend on past play of opponent(s).
- Puts zero probabilities on sub-optimal responses.

FP vs. Smoothed FP vs. Exponentiated RM

Fictitious play Plays *best responses*.

- Does depend on past play of opponent(s).
- Puts zero probabilities on sub-optimal responses.

Smoothed fictitious play May play *sub-optimal responses*, e.g., proportional to their estimated payoffs.

FP vs. Smoothed FP vs. Exponentiated RM

Fictitious play Plays *best responses*.

- Does depend on past play of opponent(s).
- Puts zero probabilities on sub-optimal responses.

Smoothed fictitious play May play *sub-optimal responses*, e.g., proportional to their estimated payoffs.

- Does depend on past play of opponent(s).

FP vs. Smoothed FP vs. Exponentiated RM

Fictitious play Plays *best responses*.

- Does depend on past play of opponent(s).
- Puts zero probabilities on sub-optimal responses.

Smoothed fictitious play May play *sub-optimal responses*, e.g., proportional to their estimated payoffs.

- Does depend on past play of opponent(s).
- Puts non-zero probabilities on sub-optimal responses.

FP vs. Smoothed FP vs. Exponentiated RM

Fictitious play Plays *best responses*.

- Does depend on past play of opponent(s).
- Puts zero probabilities on sub-optimal responses.

Smoothed fictitious play May play *sub-optimal responses*, e.g., proportional to their estimated payoffs.

- Does depend on past play of opponent(s).
- Puts non-zero probabilities on sub-optimal responses.

- Approaches fictitious play when $\gamma \downarrow 0$ (PY, p. 84).

FP vs. Smoothed FP vs. Exponentiated RM

Fictitious play Plays *best responses*.

- Does depend on past play of opponent(s).
- Puts zero probabilities on sub-optimal responses.

Smoothed fictitious play May play *sub-optimal responses*, e.g., proportional to their estimated payoffs.

- Does depend on past play of opponent(s).
- Puts non-zero probabilities on sub-optimal responses.
- Approaches fictitious play when $\gamma \downarrow 0$ (PY, p. 84).

Exponentiated regret matching Plays *regret suboptimally*, i.e., proportional to a power of positive regret.

FP vs. Smoothed FP vs. Exponentiated RM

Fictitious play Plays *best responses*.

- Does depend on past play of opponent(s).
- Puts zero probabilities on sub-optimal responses.

Smoothed fictitious play May play *sub-optimal responses*, e.g., proportional to their estimated payoffs.

- Does depend on past play of opponent(s).
- Puts non-zero probabilities on sub-optimal responses.
- Approaches fictitious play when $\gamma \downarrow 0$ (PY, p. 84).

Exponentiated regret matching Plays *regret suboptimally*, i.e., proportional to a power of positive regret.

- Does depend on own past payoffs.

FP vs. Smoothed FP vs. Exponentiated RM

Fictitious play Plays *best responses*.

- Does depend on past play of opponent(s).
- Puts zero probabilities on sub-optimal responses.

Smoothed fictitious play May play *sub-optimal responses*, e.g., proportional to their estimated payoffs.

- Does depend on past play of opponent(s).
- Puts non-zero probabilities on sub-optimal responses.
- Approaches fictitious play when $\gamma \downarrow 0$ (PY, p. 84).

Exponentiated regret matching Plays *regret suboptimally*, i.e., proportional to a power of positive regret.

- Does depend on own past payoffs.
- Puts non-zero probabilities on sub-optimal responses.

FP vs. Smoothed FP vs. Exponentiated RM

Fictitious play Plays *best responses*.

- Does depend on past play of opponent(s).
- Puts zero probabilities on sub-optimal responses.

Smoothed fictitious play May play *sub-optimal responses*, e.g., proportional to their estimated payoffs.

- Does depend on past play of opponent(s).
- Puts non-zero probabilities on sub-optimal responses.
- Approaches fictitious play when $\gamma \downarrow 0$ (PY, p. 84).

Exponentiated regret matching Plays *regret suboptimally*, i.e., proportional to a power of positive regret.

- Does depend on own past payoffs.
- Puts non-zero probabilities on sub-optimal responses.
- Approaches fictitious play when exponent $a \to \infty$ (PY, p. 84).

FP vs. Smoothed FP vs. Exponentiated RM

FP vs. Smoothed FP vs. Exponentiated RM

	FP	Smoothed FP	Exponentiated RM

FP vs. Smoothed FP vs. Exponentiated RM

	FP	Smoothed FP	Exponentiated RM
Depends on past play of opponents	$\surd$	$\surd$	$-$

FP vs. Smoothed FP vs. Exponentiated RM

	FP	**Smoothed FP**	**Exponentiated RM**
Depends on past play of opponents	$\surd$	$\surd$	$-$
Depends on own past payoffs	$-$	$-$	$\surd$

FP vs. Smoothed FP vs. Exponentiated RM

	FP	Smoothed FP	Exponentiated RM
Depends on past play of opponents	$\surd$	$\surd$	—
Depends on own past payoffs	—	—	$\surd$
Puts zero probabilities on sub-optimal responses	$\surd$	—	—

FP vs. Smoothed FP vs. Exponentiated RM

	FP	**Smoothed FP**	**Exponentiated RM**
Depends on past play of opponents	$\surd$	$\surd$	–
Depends on own past payoffs	–	–	$\surd$
Puts zero probabilities on sub-optimal responses	$\surd$	–	–
Best response	$\surd$	when smoothing parameter $\gamma \downarrow 0$	when exponent $a \to \infty$

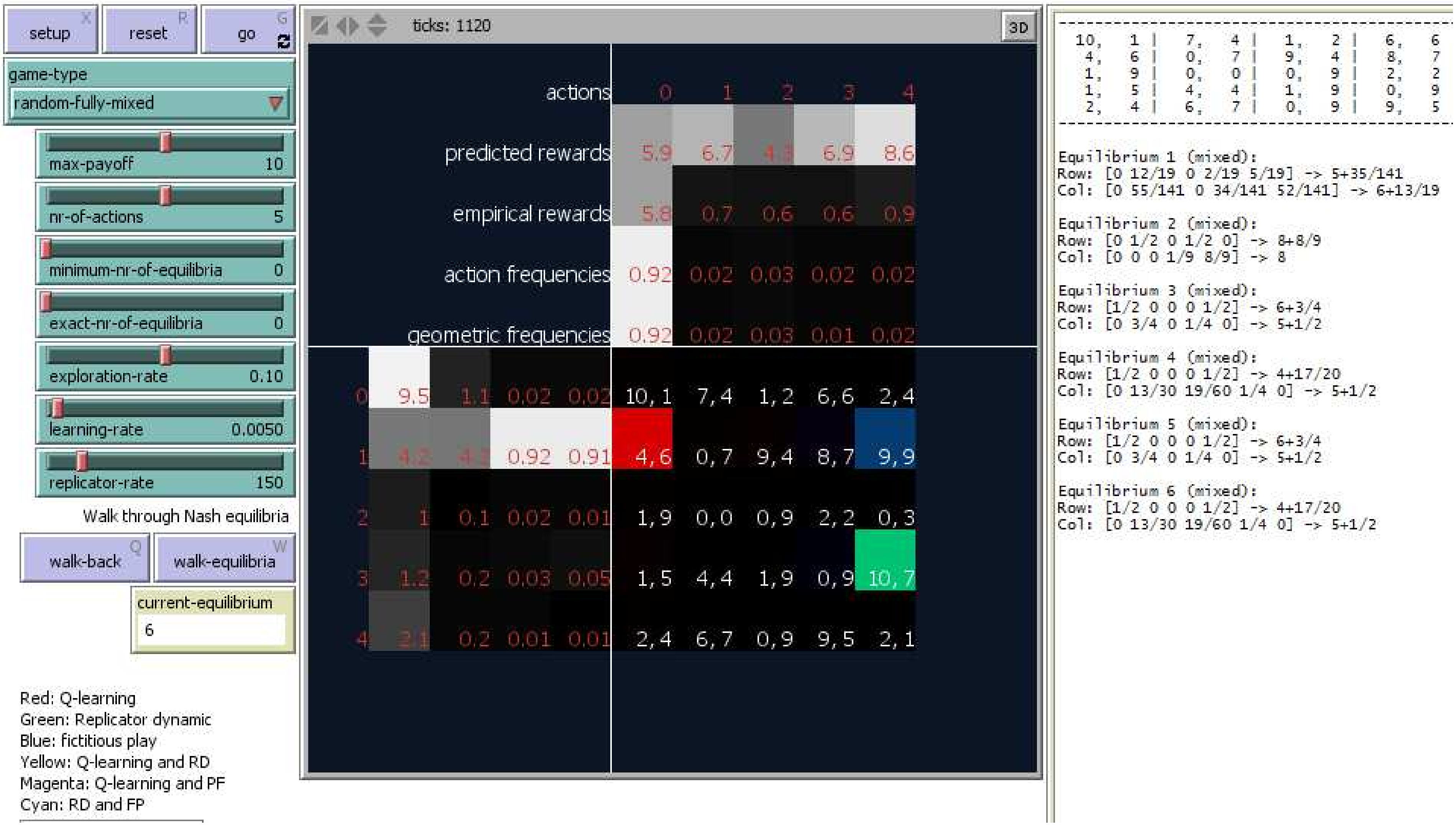
FP vs. Smoothed FP vs. Exponentiated RM

	FP	Smoothed FP	Exponentiated RM
Depends on past play of opponents	$\surd$	$\surd$	–
Depends on own past payoffs	–	–	$\surd$
Puts zero probabilities on sub-optimal responses	$\surd$	–	–
Best response	$\surd$	when smoothing parameter $\gamma \downarrow 0$	when exponent $a \to \infty$
Individual no-regret	–	Within $\epsilon > 0$, almost always (PY, p. 82)	Exact, almost always (PY, p. 60)

FP vs. Smoothed FP vs. Exponentiated RM

	FP	Smoothed FP	Exponentiated RM
Depends on past play of opponents	$\surd$	$\surd$	$-$
Depends on own past payoffs	$-$	$-$	$\surd$
Puts zero probabilities on sub-optimal responses	$\surd$	$-$	$-$
Best response	$\surd$	when smoothing parameter $\gamma \downarrow 0$	when exponent $a \to \infty$
Individual no-regret	$-$	Within $\epsilon > 0$, almost always (PY, p. 82)	Exact, almost always (PY, p. 60)
Collective convergence to coarse correlated equilibria	$-$	Within $\epsilon > 0$, almost always (PY, p. 83)	Exact, almost always (PY, p. 60)

Fictitious play compared to other algorithms



Part III:
Finite memory and inertia

Finite memory: motivation

Finite memory: motivation

- In their basic version, most learning rules rely on the entire history of play.

Finite memory: motivation

- In their basic version, most learning rules rely on the entire history of play.
- People, as well as computers, have a finite memory.

Finite memory: motivation

- In their basic version, most learning rules rely on the entire history of play.
- People, as well as computers, have a finite memory. (On the other hand, for average or discounted payoffs this is no problem.)

Finite memory: motivation

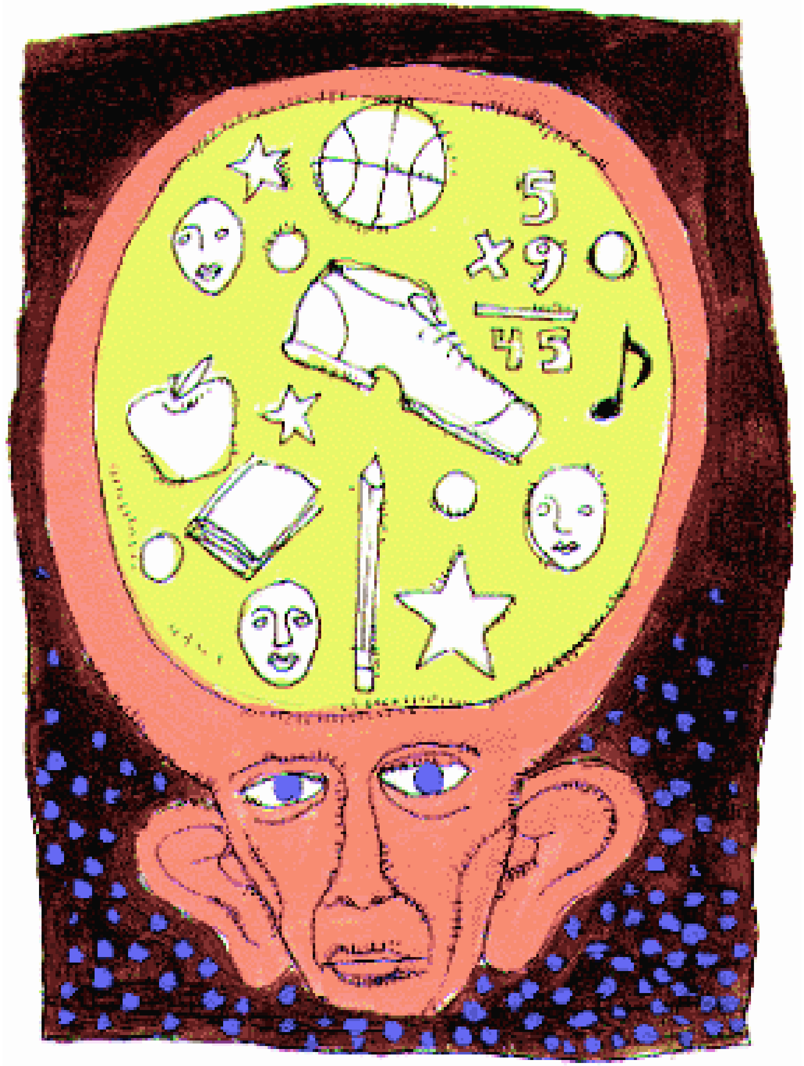
- In their basic version, most learning rules rely on the entire history of play.
- People, as well as computers, have a finite memory. (On the other hand, for average or discounted payoffs this is no problem.)
- Nevertheless: experiences in the distant past are apt to be less relevant than more recent ones.

Finite memory: motivation

- In their basic version, most learning rules rely on the entire history of play.
- People, as well as computers, have a finite memory. (On the other hand, for average or discounted payoffs this is no problem.)
- Nevertheless: experiences in the distant past are apt to be less relevant than more recent ones.
- Idea: let players have a finite memory of **m** rounds.

Finite memory: motivation

- In their basic version, most learning rules rely on the entire history of play.
- People, as well as computers, have a finite memory. (On the other hand, for average or discounted payoffs this is no problem.)
- Nevertheless: experiences in the distant past are apt to be less relevant than more recent ones.
- Idea: let players have a finite memory of m rounds.



Inertia: motivation

Inertia: motivation

- When players' strategies are constantly re-evaluated, discontinuities in behaviour are likely to occur.

Inertia: motivation

- When players' strategies are constantly re-evaluated, discontinuities in behaviour are likely to occur.

  Example: the asymmetric coordination game.

Inertia: motivation

- When players' strategies are constantly re-evaluated, discontinuities in behaviour are likely to occur.

  Example: the asymmetric coordination game.

- Discontinuities in behaviour are less likely to lead to equilibria of any sort.

Inertia: motivation

- When players' strategies are constantly re-evaluated, discontinuities in behaviour are likely to occur.

  Example: the asymmetric coordination game.

- Discontinuities in behaviour are less likely to lead to equilibria of any sort.

- Idea: let players play the same action as in the previous round with probability λ.

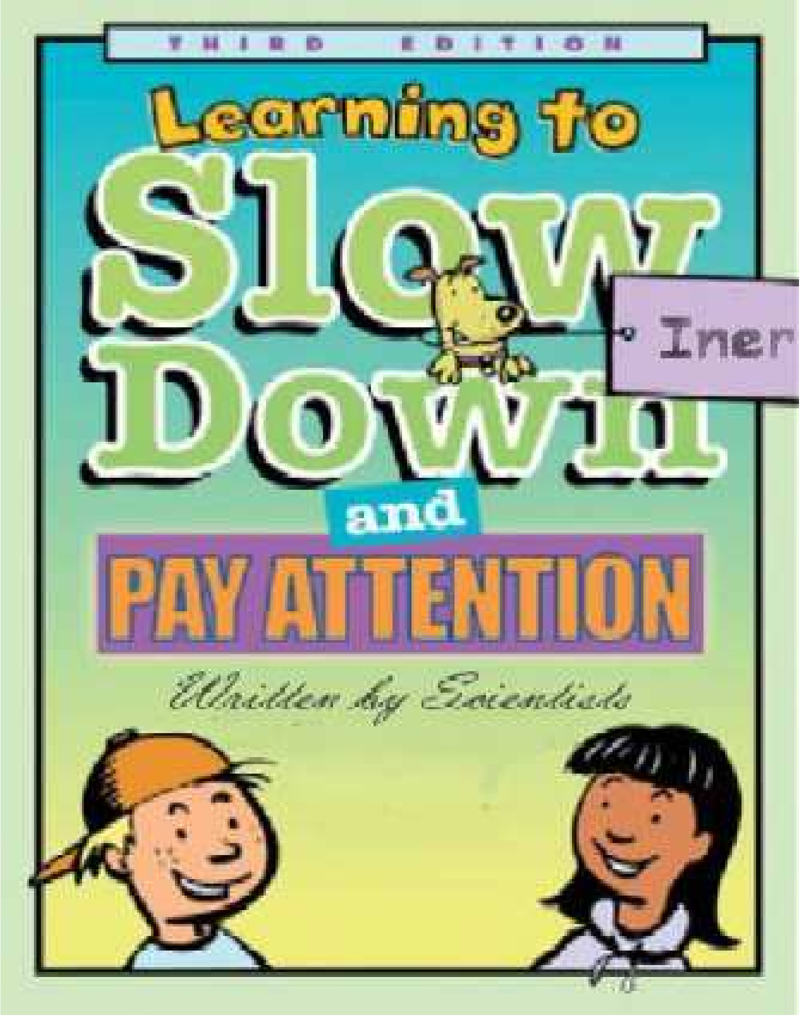
Inertia: motivation

- When players' strategies are constantly re-evaluated, discontinuities in behaviour are likely to occur.

  Example: the asymmetric coordination game.

- Discontinuities in behaviour are less likely to lead to equilibria of any sort.

- Idea: let players play the same action as in the previous round with probability λ.



Weakly acyclic games

Weakly acyclic games

- Game G with action space X.

Weakly acyclic games

- Game G with action space X.
- $G' = (V, E)$ where $V = X$ and

$$E = \{ (x, y) \mid \text{for some } i : y_{-i} = x_{-i} \text{ and } u_i(y_i, y_{-i}) > u_i(x_i, x_{-i}) \}.$$

Weakly acyclic games

- Game G with action space X.
- $G' = (V, E)$ where $V = X$ and

$$E = \{ (x, y) \mid \text{for some } i : y_{-i} = x_{-i} \text{ and } u_i(y_i, y_{-i}) > u_i(x_i, x_{-i}) \}.$$

- For all $x \in X$: x is a sink iff x is a pure Nash equilibrium.

Weakly acyclic games

- Game G with action space X.
- $G' = (V, E)$ where $V = X$ and

$$E = \{ (x, y) \mid \text{for some } i : \\ y_{-i} = x_{-i} \text{ and } u_i(y_i, y_{-i}) > \\ u_i(x_i, x_{-i}) \}.$$

- For all $x \in X$: x is a sink iff x is a pure Nash equilibrium.
- G is said to be weakly acyclic under better replies if every node is connected to a sink.

Weakly acyclic games

- Game G with action space X.
- $G' = (V, E)$ where $V = X$ and

$$E = \{ (x, y) \mid \text{for some } i : \\ y_{-i} = x_{-i} \text{ and } u_i(y_i, y_{-i}) > \\ u_i(x_i, x_{-i}) \}.$$

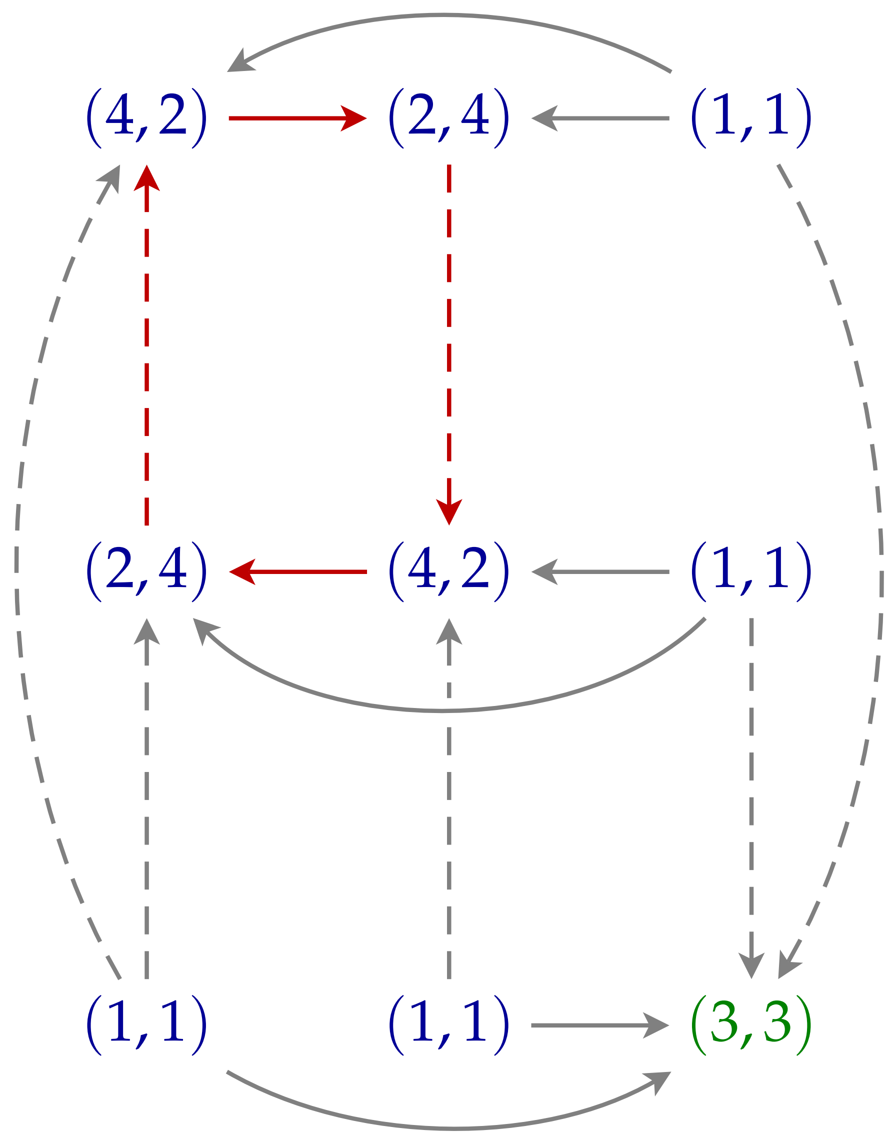
- For all $x \in X$: x is a sink iff x is a pure Nash equilibrium.
- G is said to be weakly acyclic under better replies if every node is connected to a sink.
- WAuBR $\Rightarrow \exists$ pure Nash equilibrium.

Weakly acyclic games

- Game G with action space X.
- $G' = (V, E)$ where $V = X$ and

$$E = \{ (x, y) \mid \text{for some } i : y_{-i} = x_{-i} \text{ and } u_i(y_i, y_{-i}) > u_i(x_i, x_{-i}) \}.$$

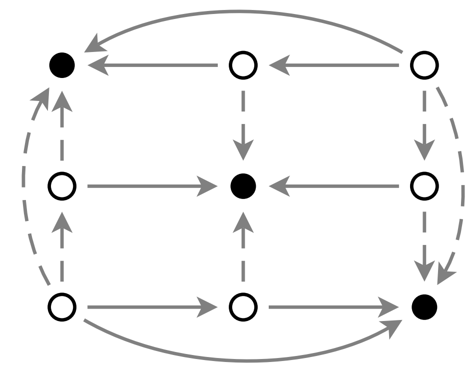
- For all $x \in X$: x is a sink iff x is a pure Nash equilibrium.
- G is said to be weakly acyclic under better replies if every node is connected to a sink.
- WAuBR $\Rightarrow \exists$ pure Nash equilibrium.



Examples of weakly acyclic games

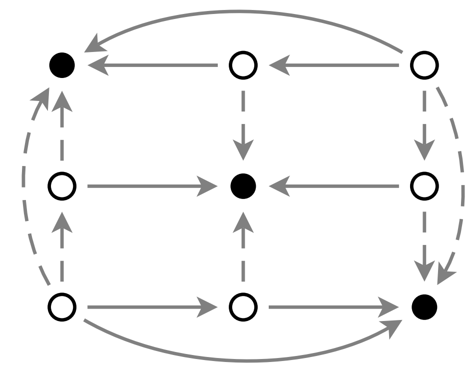
Examples of weakly acyclic games

Coordination games Two-person games with identical actions for all players, where best responses are formed by the diagonal of the joint action space.



Examples of weakly acyclic games

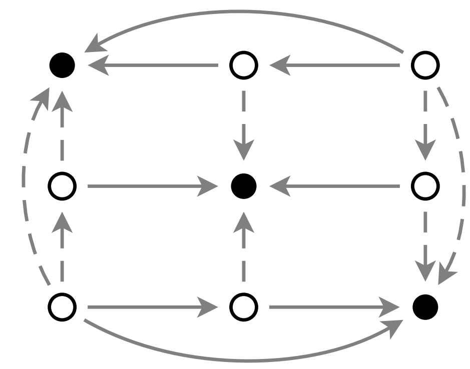
Coordination games Two-person games with identical actions for all players, where best responses are formed by the diagonal of the joint action space.



Potential games (Monderer and Shapley, 1996).

There is a function $\rho : X \to R$, called the potential,

Examples of weakly acyclic games



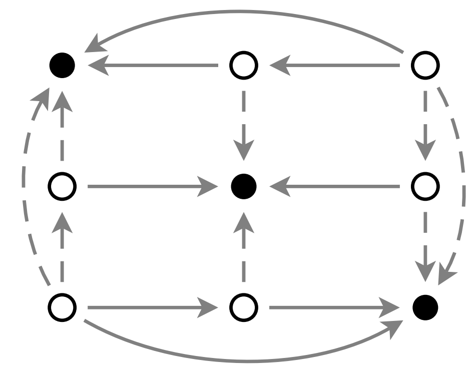
Coordination games Two-person games with identical actions for all players, where best responses are formed by the diagonal of the joint action space.

Potential games (Monderer and Shapley, 1996).

There is a function $\rho : X \to R$, called the potential, such that for every player i

Examples of weakly acyclic games

Coordination games Two-person games with identical actions for all players, where best responses are formed by the diagonal of the joint action space.

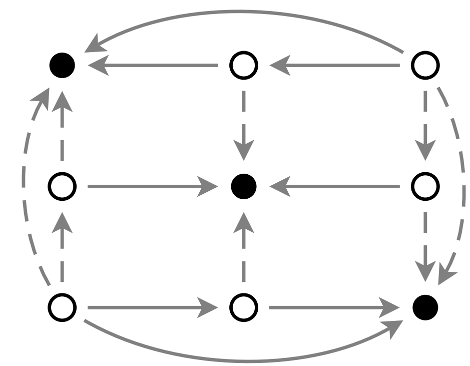


Potential games (Monderer and Shapley, 1996).

There is a function $\rho : X \to R$, called the potential, such that for every player i and every action profile $x, y \in X$:

Examples of weakly acyclic games

Coordination games Two-person games with identical actions for all players, where best responses are formed by the diagonal of the joint action space.



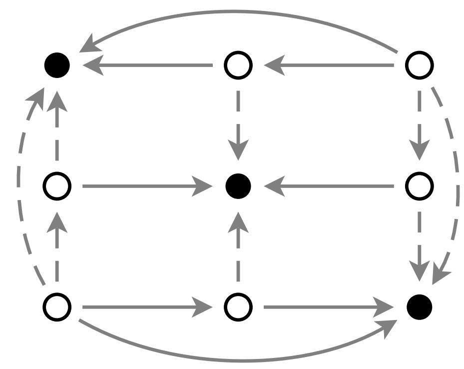
Potential games (Monderer and Shapley, 1996).

There is a function $\rho : X \to R$, called the potential, such that for every player i and every action profile $x, y \in X$:

$$y_{-i} = x_{-i} \;\Rightarrow\; u_i(y_i, y_{-i}) - u_i(x_i, x_{-i}) = \rho(y) - \rho(x)$$

Examples of weakly acyclic games

Coordination games Two-person games with identical actions for all players, where best responses are formed by the diagonal of the joint action space.



Potential games (Monderer and Shapley, 1996).

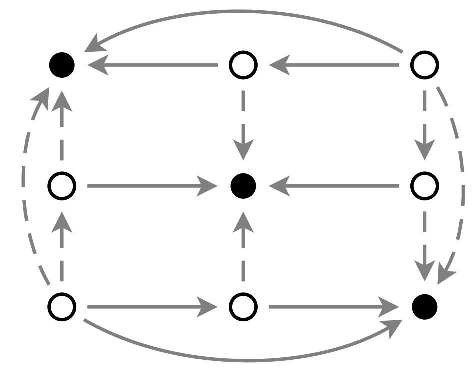
There is a function $\rho : X \to R$, called the potential, such that for every player i and every action profile $x, y \in X$:

$$y_{-i} = x_{-i} \;\Rightarrow\; u_i(y_i, y_{-i}) - u_i(x_i, x_{-i}) = \rho(y) - \rho(x)$$

Example: congestion games.

Examples of weakly acyclic games

Coordination games Two-person games with identical actions for all players, where best responses are formed by the diagonal of the joint action space.



Potential games (Monderer and Shapley, 1996).

There is a function $\rho : X \to R$, called the potential, such that for every player i and every action profile $x, y \in X$:

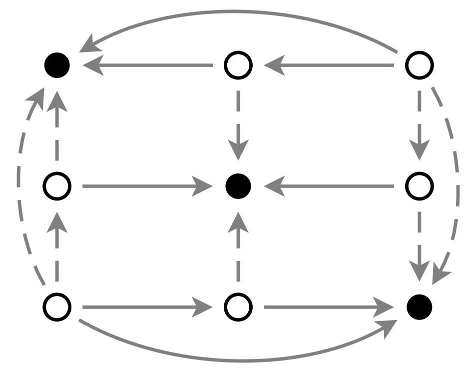
$$y_{-i} = x_{-i} \;\Rightarrow\; u_i(y_i, y_{-i}) - u_i(x_i, x_{-i}) = \rho(y) - \rho(x)$$

Example: congestion games.

true : The potential function increases along every path.

Examples of weakly acyclic games

Coordination games Two-person games with identical actions for all players, where best responses are formed by the diagonal of the joint action space.



Potential games (Monderer and Shapley, 1996).

There is a function $\rho : X \to R$, called the potential, such that for every player i and every action profile $x, y \in X$:

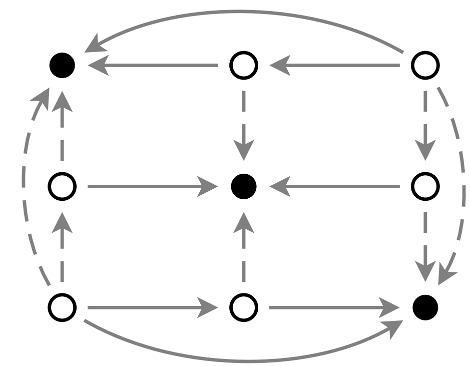
$$y_{-i} = x_{-i} \;\Rightarrow\; u_i(y_i, y_{-i}) - u_i(x_i, x_{-i}) = \rho(y) - \rho(x)$$

Example: congestion games.

true : The potential function increases along every path.
$\Rightarrow$: Paths cannot cycle.

Examples of weakly acyclic games

Coordination games Two-person games with identical actions for all players, where best responses are formed by the diagonal of the joint action space.



Potential games (Monderer and Shapley, 1996).

There is a function $\rho : X \to R$, called the potential, such that for every player i and every action profile $x, y \in X$:

$$y_{-i} = x_{-i} \Rightarrow u_i(y_i, y_{-i}) - u_i(x_i, x_{-i}) = \rho(y) - \rho(x)$$

Example: congestion games.

true : The potential function increases along every path.
$\Rightarrow$: Paths cannot cycle.
$\Rightarrow$: In finite graphs, paths must end.

Weakly acyclic games under finite memory and inertia

Theorem. *Let G be a finite weakly acyclic n-person game. Every better-reply process with finite memory and inertia converges to a pure NE of G.*

Weakly acyclic games under finite memory and inertia

Theorem. *Let G be a finite weakly acyclic n-person game. Every better-reply process with finite memory and inertia converges to a pure NE of G.*

Proof. (Outline.)

Weakly acyclic games under finite memory and inertia

Theorem. *Let G be a finite weakly acyclic n-person game. Every better-reply process with finite memory and inertia converges to a pure NE of G.*

Proof. (Outline.)

1. Let the state space Z be $X^{\mathbf{m}}$.

Weakly acyclic games under finite memory and inertia

Theorem. *Let G be a finite weakly acyclic n-person game. Every better-reply process with finite memory and inertia converges to a pure NE of G.*

Proof. (Outline.)

1. Let the state space Z be $X^{\mathbf{m}}$.

2. A state $\bar{x} \in X^{\mathbf{m}}$ is called homogeneous if it consists of identical action profiles x. Such a state is denoted by $\langle x \rangle$.

Weakly acyclic games under finite memory and inertia

Theorem. *Let G be a finite weakly acyclic n-person game. Every better-reply process with finite memory and inertia converges to a pure NE of G.*

Proof. (Outline.)

1. Let the state space Z be $X^{\mathbf{m}}$.

2. A state $\bar{x} \in X^{\mathbf{m}}$ is called homogeneous if it consists of identical action profiles x. Such a state is denoted by $\langle x \rangle$.

   $Z^* =_{Def}$
   $\{$ homogeneous states $\}$.

Weakly acyclic games under finite memory and inertia

Theorem. *Let G be a finite weakly acyclic n-person game. Every better-reply process with finite memory and inertia converges to a pure NE of G.*

Proof. (Outline.)

1. Let the state space Z be $X^{\mathbf{m}}$.

2. A state $\bar{x} \in X^{\mathbf{m}}$ is called homogeneous if it consists of identical action profiles x. Such a state is denoted by $\langle x \rangle$.

   $Z^* =_{Def}$
   $\{$ homogeneous states $\}$.

3. It can be shown that the process hits Z^* infinitely often.

Weakly acyclic games under finite memory and inertia

Theorem. *Let G be a finite weakly acyclic n-person game. Every better-reply process with finite memory and inertia converges to a pure NE of G.*

Proof. (Outline.)

1. Let the state space Z be $X^{\mathbf{m}}$.

2. A state $\bar{x} \in X^{\mathbf{m}}$ is called homogeneous if it consists of identical action profiles x. Such a state is denoted by $\langle x \rangle$.

   $Z^* =_{Def}$ { homogeneous states }.

3. It can be shown that the process hits Z^* infinitely often.

4. It can be shown that the overall probability to play any action is bounded away from zero.

Weakly acyclic games under finite memory and inertia

Theorem. *Let G be a finite weakly acyclic n-person game. Every better-reply process with finite memory and inertia converges to a pure NE of G.*

Proof. (Outline.)

1. Let the state space Z be $X^{\mathbf{m}}$.
2. A state $\bar{x} \in X^{\mathbf{m}}$ is called homogeneous if it consists of identical action profiles x. Such a state is denoted by $\langle x \rangle$.

   $Z^* =_{Def}$ $\{$ homogeneous states $\}$.
3. It can be shown that the process hits Z^* infinitely often.
4. It can be shown that the overall probability to play any action is bounded away from zero.
5. It can be shown that

   $$\text{Absorbing} \overset{\subseteq}{\supseteq} Z^* \cap \text{ Pure Nash.}$$

Weakly acyclic games under finite memory and inertia

Theorem. *Let G be a finite weakly acyclic n-person game. Every better-reply process with finite memory and inertia converges to a pure NE of G.*

Proof. (Outline.)

1. Let the state space Z be $X^{\mathbf{m}}$.
2. A state $\bar{x} \in X^{\mathbf{m}}$ is called homogeneous if it consists of identical action profiles x. Such a state is denoted by $\langle x \rangle$.

   $Z^* =_{Def}$ $\{$ homogeneous states $\}$.
3. It can be shown that the process hits Z^* infinitely often.
4. It can be shown that the overall probability to play any action is bounded away from zero.
5. It can be shown that

   $$\text{Absorbing} \begin{smallmatrix}\subseteq\\ \supseteq\end{smallmatrix} Z^* \cap \text{ Pure Nash.}$$
6. It can be shown that, due to weak acyclicity, inertia, and (4), the process eventually lands in an absorbing state which, due to (5), is a repeated pure Nash equilibrium. $\square$

First claim: process hits Z^* infinitely often

First claim: process hits Z^* infinitely often

Let inertia be determined by $\lambda > 0$.

$$\Pr(\text{all players play their previous action}) = \lambda^n.$$

First claim: process hits Z^* infinitely often

Let inertia be determined by $\lambda > 0$.

$$\Pr(\text{all players play their previous action}) = \lambda^n.$$

Hence,

$\Pr(\text{all pl. play their previous action during } \mathbf{m} \text{ subsequent rounds}) = \lambda^{n\mathbf{m}}$.

First claim: process hits Z^* infinitely often

Let inertia be determined by $\lambda > 0$.

$$\Pr(\text{all players play their previous action}) = \lambda^n.$$

Hence,

$\Pr(\text{all pl. play their previous action during } \mathbf{m} \text{ subsequent rounds}) = \lambda^{n\mathbf{m}}$.

If all players play their previous action during **m** subsequent rounds, then the process arrives at a homogeneous state.

First claim: process hits Z^* infinitely often

Let inertia be determined by $\lambda > 0$.

$$\Pr(\text{all players play their previous action}) = \lambda^n.$$

Hence,

$$\Pr(\text{all pl. play their previous action during } \mathbf{m} \text{ subsequent rounds}) = \lambda^{n\mathbf{m}}.$$

If all players play their previous action during **m** subsequent rounds, then the process arrives at a homogeneous state. Therefore, for all t,

$$\Pr(\text{process arrives at a homogeneous state in round } t + \mathbf{m}) \geq \lambda^{n\mathbf{m}}.$$

First claim: process hits Z^* infinitely often

Let inertia be determined by $\lambda > 0$.

$$\Pr(\text{all players play their previous action}) = \lambda^n.$$

Hence,

$$\Pr(\text{all pl. play their previous action during } \mathbf{m} \text{ subsequent rounds}) = \lambda^{n\mathbf{m}}.$$

If all players play their previous action during **m** subsequent rounds, then the process arrives at a homogeneous state. Therefore, for all t,

$$\Pr(\text{process arrives at a homogeneous state in round } t + \mathbf{m}) \geq \lambda^{n\mathbf{m}}.$$

Infinitely many disjoint histories of length **m** occur, hence infinitely many *independent* events "homogeneous at $t + \mathbf{m}$" occur.

First claim: process hits Z^* infinitely often

Let inertia be determined by $\lambda > 0$.

$$\Pr(\text{all players play their previous action}) = \lambda^n.$$

Hence,

$$\Pr(\text{all pl. play their previous action during } \mathbf{m} \text{ subsequent rounds}) = \lambda^{n\mathbf{m}}.$$

If all players play their previous action during **m** subsequent rounds, then the process arrives at a homogeneous state. Therefore, for all t,

$$\Pr(\text{process arrives at a homogeneous state in round } t + \mathbf{m}) \geq \lambda^{n\mathbf{m}}.$$

Infinitely many disjoint histories of length **m** occur, hence infinitely many *independent* events "homogeneous at $t + \mathbf{m}$" occur. Apply the (second) Borel-Cantelli lemma: if $\{E_n\}_n$ are independent events, and $\sum_{n=1}^{\infty} \Pr(E_n)$ is unbounded, then $\Pr(\text{ an infinite number of } E_n \text{ 's occur }) = 1$. □

Second claim: probability to play any action is positive

A better-reply learning method from states (finite histories) to strategies (probability distributions on actions)

$$\gamma_i : Z \to \Delta(X_i)$$

possesses two important properties:

Second claim: probability to play any action is positive

A better-reply learning method from states (finite histories) to strategies (probability distributions on actions)

$$\gamma_i : Z \rightarrow \Delta(X_i)$$

possesses two important properties:

1. The formula itself is deterministic.

Second claim: probability to play any action is positive

A better-reply learning method from states (finite histories) to strategies (probability distributions on actions)

$$\gamma_i : Z \to \Delta(X_i)$$

possesses two important properties:

1. The formula itself is deterministic.
2. Every action is played with positive probability.

Second claim: probability to play any action is positive

A better-reply learning method from states (finite histories) to strategies (probability distributions on actions)

$$\gamma_i : Z \to \Delta(X_i)$$

possesses two important properties:

1. The formula itself is deterministic.
2. Every action is played with positive probability.

Therefore, if

$$\gamma_i =_{Def} \inf\{\gamma_i(z)(x_i) \mid z \in Z,\ x_i \in X_i\}.$$

Second claim: probability to play any action is positive

A better-reply learning method from states (finite histories) to strategies (probability distributions on actions)

$$\gamma_i : Z \to \Delta(X_i)$$

possesses two important properties:

1. The formula itself is deterministic.
2. Every action is played with positive probability.

Therefore, if

$$\gamma_i =_{Def} \inf\{\gamma_i(z)(x_i) \mid z \in Z,\, x_i \in X_i\}.$$

then $\gamma_i > 0$ since all $\gamma_i(z)(x_i) > 0$ and Z and X_i are finite.

Second claim: probability to play any action is positive

A better-reply learning method from states (finite histories) to strategies (probability distributions on actions)

$$\gamma_i : Z \to \Delta(X_i)$$

possesses two important properties:

1. The formula itself is deterministic.
2. Every action is played with positive probability.

Therefore, if

$$\gamma_i =_{Def} \inf\{\gamma_i(z)(x_i) \mid z \in Z,\, x_i \in X_i\}.$$

then $\gamma_i > 0$ since all $\gamma_i(z)(x_i) > 0$ and Z and X_i are finite.

So if

$$\gamma =_{Def} \min\{\gamma_i \mid 1 \leq i \leq n\}.$$

Second claim: probability to play any action is positive

A better-reply learning method from states (finite histories) to strategies (probability distributions on actions)

$$\gamma_i : Z \to \Delta(X_i)$$

possesses two important properties:

1. The formula itself is deterministic.

2. Every action is played with positive probability.

Therefore, if

$$\gamma_i =_{Def} \inf\{\gamma_i(z)(x_i) \mid z \in Z, x_i \in X_i\}.$$

then $\gamma_i > 0$ since all $\gamma_i(z)(x_i) > 0$ and Z and X_i are finite.

So if

$$\gamma =_{Def} \min\{\gamma_i \mid 1 \leq i \leq n\}.$$

then $\gamma > 0$. $\square$

Final claim: probability to reach a sink from $Z^* > 0$

Final claim: probability to reach a sink from $\mathbf{Z}^* > 0$

Suppose the process is in $\langle x \rangle$.

Final claim: probability to reach a sink from $\mathbf{Z}^* > 0$

Suppose the process is in $\langle x \rangle$.

1. If x is pure Nash, we are done, because response functions are deterministic better replies.

Final claim: probability to reach a sink from $Z^* > 0$

Suppose the process is in $\langle x \rangle$.

1. If x is pure Nash, we are done, because response functions are deterministic better replies.

2. If x is not pure Nash, there must be an edge $x \rightarrow y$ in the better reply graph.

Final claim: probability to reach a sink from $Z^* > 0$

Suppose the process is in $\langle x \rangle$.

1. If x is pure Nash, we are done, because response functions are deterministic better replies.

2. If x is not pure Nash, there must be an edge $x \rightarrow y$ in the better reply graph. Suppose this edge concerns action x_i of player i.

Final claim: probability to reach a sink from $Z^* > 0$

Suppose the process is in $\langle x \rangle$.

1. If x is pure Nash, we are done, because response functions are deterministic better replies.

2. If x is not pure Nash, there must be an edge $x \rightarrow y$ in the better reply graph. Suppose this edge concerns action x_i of player i. We now know that x_i is played with probability at least γ, irrespective of player and state.

Final claim: probability to reach a sink from $Z^* > 0$

Suppose the process is in $\langle x \rangle$.

1. If x is pure Nash, we are done, because response functions are deterministic better replies.

2. If x is not pure Nash, there must be an edge $x \rightarrow y$ in the better reply graph. Suppose this edge concerns action x_i of player i. We now know that x_i is played with probability at least γ, irrespective of player and state.

Further probabilities:

Final claim: probability to reach a sink from $Z^* > 0$

Suppose the process is in $\langle x \rangle$.

1. If x is pure Nash, we are done, because response functions are deterministic better replies.

2. If x is not pure Nash, there must be an edge $x \to y$ in the better reply graph. Suppose this edge concerns action x_i of player i. We now know that x_i is played with probability at least γ, irrespective of player and state.

Further probabilities:

- All other players $j \neq i$ keep playing the same action : λ^{n-1}.

Final claim: probability to reach a sink from $Z^* > 0$

Suppose the process is in $\langle x \rangle$.

1. If x is pure Nash, we are done, because response functions are deterministic better replies.

2. If x is not pure Nash, there must be an edge $x \rightarrow y$ in the better reply graph. Suppose this edge concerns action x_i of player i. We now know that x_i is played with probability at least γ, irrespective of player and state.

Further probabilities:

- All other players $j \neq i$ keep playing the same action : λ^{n-1}.
- Edge $x \rightarrow y$ is actually traversed : $\gamma\lambda^{n-1}$.

Final claim: probability to reach a sink from $Z^* > 0$

Suppose the process is in $\langle x \rangle$.

1. If x is pure Nash, we are done, because response functions are deterministic better replies.

2. If x is not pure Nash, there must be an edge $x \rightarrow y$ in the better reply graph. Suppose this edge concerns action x_i of player i. We now know that x_i is played with probability at least γ, irrespective of player and state.

Further probabilities:

- All other players $j \neq i$ keep playing the same action : λ^{n-1}.
- Edge $x \rightarrow y$ is actually traversed : $\gamma\lambda^{n-1}$.
- Profile y is maintained for another $\mathbf{m} - 1$ rounds, so as to arrive at state $\langle y \rangle$: $\lambda^{n(\mathbf{m}-1)}$.

Final claim: probability to reach a sink from $Z^* > 0$

Suppose the process is in $\langle x \rangle$.

1. If x is pure Nash, we are done, because response functions are deterministic better replies.

2. If x is not pure Nash, there must be an edge $x \rightarrow y$ in the better reply graph. Suppose this edge concerns action x_i of player i. We now know that x_i is played with probability at least γ, irrespective of player and state.

Further probabilities:

- All other players $j \neq i$ keep playing the same action : λ^{n-1}.
- Edge $x \rightarrow y$ is actually traversed : $\gamma\lambda^{n-1}$.
- Profile y is maintained for another $\mathbf{m} - 1$ rounds, so as to arrive at state $\langle y \rangle$: $\lambda^{n(\mathbf{m}-1)}$.
- To traverse from $\langle x \rangle$ to $\langle y \rangle$: $\gamma\lambda^{n-1} \cdot \lambda^{n(\mathbf{m}-1)} = \gamma\lambda^{n\mathbf{m}-1}$.

Final claim: probability to reach a sink from $Z^* > 0$

Suppose the process is in $\langle x \rangle$.

1. If x is pure Nash, we are done, because response functions are deterministic better replies.

2. If x is not pure Nash, there must be an edge $x \to y$ in the better reply graph. Suppose this edge concerns action x_i of player i. We now know that x_i is played with probability at least γ, irrespective of player and state.

Further probabilities:

- All other players $j \neq i$ keep playing the same action : λ^{n-1}.
- Edge $x \to y$ is actually traversed : $\gamma\lambda^{n-1}$.
- Profile y is maintained for another $\mathbf{m} - 1$ rounds, so as to arrive at state $\langle y \rangle$: $\lambda^{n(\mathbf{m}-1)}$.
- To traverse from $\langle x \rangle$ to $\langle y \rangle$: $\gamma\lambda^{n-1} \cdot \lambda^{n(\mathbf{m}-1)} = \gamma\lambda^{n\mathbf{m}-1}$.
- The "image" $\langle x^{(1)} \rangle, \dots, \langle x^{(l)} \rangle$ of a better reply-path $x^{(1)}, \dots, x^{(l)}$ is followed to a sink : $\geq (\gamma\lambda^{nm-1})^L$, where L is the length of a longest better-reply path.

Final claim: probability to reach a sink from $Z^* > 0$

Suppose the process is in $\langle x \rangle$.

1. If x is pure Nash, we are done, because response functions are deterministic better replies.

2. If x is not pure Nash, there must be an edge $x \to y$ in the better reply graph. Suppose this edge concerns action x_i of player i. We now know that x_i is played with probability at least γ, irrespective of player and state.

Further probabilities:

- All other players $j \neq i$ keep playing the same action : λ^{n-1}.
- Edge $x \to y$ is actually traversed : $\gamma\lambda^{n-1}$.
- Profile y is maintained for another $\mathbf{m} - 1$ rounds, so as to arrive at state $\langle y \rangle$: $\lambda^{n(\mathbf{m}-1)}$.
- To traverse from $\langle x \rangle$ to $\langle y \rangle$: $\gamma\lambda^{n-1} \cdot \lambda^{n(\mathbf{m}-1)} = \gamma\lambda^{n\mathbf{m}-1}$.
- The "image" $\langle x^{(1)} \rangle, \dots, \langle x^{(l)} \rangle$ of a better reply-path $x^{(1)}, \dots, x^{(l)}$ is followed to a sink : $\geq (\gamma\lambda^{nm-1})^L$, where L is the length of a longest better-reply path.

Since Z^* is encountered infinitely often, the result follows. $\square$

Summary

Summary

- With fictitious play, the behaviour of opponents is modelled by (or represented by, or projected on) a mixed strategy.

Summary

- With fictitious play, the behaviour of opponents is modelled by (or represented by, or projected on) a mixed strategy.
- Fictitious play ignores sub-optimal actions.

Summary

- With fictitious play, the behaviour of opponents is modelled by (or represented by, or projected on) a mixed strategy.
- Fictitious play ignores sub-optimal actions.
- There is a family of so-called better-reply learning rules, that
  i) play sub-optimal actions, and
  ii) can be brought arbitrary close to fictitious play.

Summary

- With fictitious play, the behaviour of opponents is modelled by (or represented by, or projected on) a mixed strategy.
- Fictitious play ignores sub-optimal actions.
- There is a family of so-called better-reply learning rules, that
  i) play sub-optimal actions, and
  ii) can be brought arbitrary close to fictitious play.
- In weakly acyclic n-person games, every better-reply process with finite memory and inertia converges to a pure Nash equilibrium.

What next?

Gradient dynamics:

- Like fictitious play, players model (or assess) each other through mixed strategies.
- Strategies are not played, only maintained.
- Due to CKR (common knowledge of rationality, cf. Hargreaves Heap & Varoufakis, 2004), all models of mixed strategies are correct. (I.e., $q^{-i} = s^{-i}$, for all i.)
- Players gradually adapt their mixed strategies through hill-climbing in the payoff space.

What next?

Bayesian play:

Gradient dynamics:

- Like fictitious play, players model (or assess) each other through mixed strategies.
- Strategies are not played, only maintained.
- Due to CKR (common knowledge of rationality, cf. Hargreaves Heap & Varoufakis, 2004), all models of mixed strategies are correct. (I.e., $q^{-i} = s^{-i}$, for all i.)
- Players gradually adapt their mixed strategies through hill-climbing in the payoff space.

What next?

Bayesian play:

- With fictitious play, the behaviour of opponents is modelled by a single mixed strategy.

Gradient dynamics:

- Like fictitious play, players model (or assess) each other through mixed strategies.
- Strategies are not played, only maintained.
- Due to CKR (common knowledge of rationality, cf. Hargreaves Heap & Varoufakis, 2004), all models of mixed strategies are correct. (I.e., $q^{-i} = s^{-i}$, for all i.)
- Players gradually adapt their mixed strategies through hill-climbing in the payoff space.

What next?

Bayesian play:

- With fictitious play, the behaviour of opponents is modelled by a single mixed strategy.
- With Bayesian play, opponents are modelled by a probability distribution over (a possibly confined set of) mixed strategies.

What next?

Bayesian play:

- With fictitious play, the behaviour of opponents is modelled by a single mixed strategy.
- With Bayesian play, opponents are modelled by a probability distribution over (a possibly confined set of) mixed strategies.

Gradient dynamics:

What next?

Bayesian play:

- With fictitious play, the behaviour of opponents is modelled by a single mixed strategy.
- With Bayesian play, opponents are modelled by a probability distribution over (a possibly confined set of) mixed strategies.

Gradient dynamics:

- Like fictitious play, players model (or assess) each other through mixed strategies.

What next?

Bayesian play:

- With fictitious play, the behaviour of opponents is modelled by a single mixed strategy.
- With Bayesian play, opponents are modelled by a probability distribution over (a possibly confined set of) mixed strategies.

Gradient dynamics:

- Like fictitious play, players model (or assess) each other through mixed strategies.
- Strategies are not played, only maintained.

What next?

Bayesian play:

- With fictitious play, the behaviour of opponents is modelled by a single mixed strategy.
- With Bayesian play, opponents are modelled by a probability distribution over (a possibly confined set of) mixed strategies.

Gradient dynamics:

- Like fictitious play, players model (or assess) each other through mixed strategies.
- Strategies are not played, only maintained.
- Due to CKR (common knowledge of rationality, cf. Hargreaves Heap & Varoufakis, 2004), all models of mixed strategies are correct. (I.e., $q^{-i} = s^{-i}$, for all i.)

What next?

Bayesian play:

- With fictitious play, the behaviour of opponents is modelled by a single mixed strategy.
- With Bayesian play, opponents are modelled by a probability distribution over (a possibly confined set of) mixed strategies.

Gradient dynamics:

- Like fictitious play, players model (or assess) each other through mixed strategies.
- Strategies are not played, only maintained.
- Due to CKR (common knowledge of rationality, cf. Hargreaves Heap & Varoufakis, 2004), all models of mixed strategies are correct. (I.e., $q^{-i} = s^{-i}$, for all i.)
- Players gradually adapt their mixed strategies through hill-climbing in the payoff space.