6.207/14.15: Networks
Lectures 13 and 14: Evolution and Learning in Games

Daron Acemoglu and Asu Ozdaglar
MIT

October 26 and 28, 2009

Outline

- Myopic and Rule of Thumb Behavior
- Evolution
- Evolutionarily Stable Strategies
- Replicator Dynamics
- Learning in Games
- Fictitious Play
- Convergence of Fictitious Play in Potential Games
- Rule of Thumb Behavior and Nash Equilibrium.

- Reading:
- Osborne, Chapters 13.
- EK, Chapter 7.

Motivation

- Do people play Nash equilibrium?
- In class, in the context of the k-beauty game, we saw that even very smart MIT students do not play the unique Nash equilibrium (or the unique strategy profile surviving iterated elimination of strictly dominated strategies).
- Why?
  - Either because in new situations, it is often quite complex to work out what is "best".
  - Or more likely, because, again in new situations, individuals are uncertain about how others will play the game.
- If we played the k-beauty game several more times, behavior would have approached or in fact reached the Nash equilibrium prediction.

Motivation (continued)

- This reasoning suggests the following:
- Perhaps people behave using simple rules of thumb; these are somewhat "myopic," in the sense that they do not involve full computation of optimal strategies for others and for oneself.
- But they are also "flexible" rules of thumb in the sense that they adapt and respond to situations, including to the (actual) behavior of other players.
- What are the implications of this type of adaptive behavior?
- Two different and complementary approaches:
    1. Evolutionary game theory.
    2. Learning in games.

Evolution and Game Theory

- The theory of evolution goes back to Darwin's classic, *The Origins of Species* (and to Wallace).
- Darwin focused mostly on evolution and adaptation of an organism to the environment in which it was situated. But in *The Descent of Man*, in the context of sexual selection, he anticipated many of the ideas of evolutionary game theory.
- Evolutionary game theory was introduced by John Maynard Smith in *Evolution and the Theory of Games*, and in his seminal papers, Maynard Smith (1972) "Game Theory and the Evolution of Fighting" and Maynard Smith and Price (1973) "The Logic of Animal Conflict".
- The theory was formulated for understanding the behavior of animals in game-theoretic situations (to a game theorist, **all** situations). But it can equally well be applied to modeling "myopic behavior" for more complex organisms—such as humans.

Evolution in Strategies

- In its simplest form the story goes like this: each organism is born programmed to play a particular strategy.
- The game is the **game of life**—with payoffs given as fitness. If the organism is successful, it has greater fitness and more offspring, also programmed to play in the same way. If it is unsuccessful, it likely dies without offspring.
- **Mutations** imply that some of these offspring will randomly play any one of the feasible strategies.
- This situation can then be modeled in two different ways:
  1. By defining a concept of equilibrium appropriate for this evolutionary "competition". The concept that Maynard Smith proposed is evolutionary stability.
  2. By defining the dynamics of evolution more explicitly through replicator dynamics.
- Note: many other uses of "evolutionary" ideas in economics.

The Setting

- Consider a large population of agents (organisms, animals, humans).
- At each instant, each agent is **randomly matched** with one other agent, and they play a symmetric strategic form game. The payoffs of the game are their **fitness** level.
- Each agent is programmed (committed to) to playing a given strategy.
- Strategies that have higher payoffs expand and those that have lower payoffs contract.

A Reminder and A New Concept

Definition

(Nash equilibrium) *A (pure strategy) Nash Equilibrium of a strategic form game* $\langle \mathcal{I}, (S_i)_{i \in \mathcal{I}}, (u_i)_{i \in \mathcal{I}} \rangle$ *is a strategy profile* $s^* \in S$ *such that for all* $i \in \mathcal{I}$

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*) \qquad \text{for all } s_i \in S_i.$$

Definition

(Strict Nash equilibrium) *A strict Nash Equilibrium of a strategic form game* $\langle \mathcal{I}, (S_i)_{i \in \mathcal{I}}, (u_i)_{i \in \mathcal{I}} \rangle$ *is a strategy profile* $s^* \in S$ *such that for all* $i \in \mathcal{I}$

$$u_i(s_i^*, s_{-i}^*) > u_i(s_i, s_{-i}^*) \qquad \text{for all } s_i \in S_i.$$

- Clearly, a strict Nash equilibrium may not exist.
- Also, a strict Nash equilibrium is necessarily a pure strategy equilibrium (why?).

The Canonical Game

- The canonical game used in much of evolutionarily game theory to motivate the main ideas is the Hawk-Dove game:

Player 1/Player 2	Hawk	Dove
Hawk	$\left(\frac{1}{2}(v-c), \frac{1}{2}(v-c)\right)$	$(v, 0)$
Dove	$(0, v)$	$\left(\frac{1}{2}v, \frac{1}{2}v\right)$

- Interpretation:
  - There is a resource of value v to be shared. If a player plays "Hawk," it is aggressive and will try to take the whole resource for itself. If the other player is playing "Dove," it will succeed in doing so. If both players are playing "Hawk," then they fight and they share the resource but lose c in the process. If they are both playing "Dove," then they just share the resource.
- Interpret the payoffs as corresponding to **fitness**—e.g., greater consumption of resources leads to more offspring.

The Canonical Game (continued)

- Depending on the value of c relative to v, there are different types of equilibria.
  - If $v > c$, then there is a unique **strict** Nash equilibrium, which is (Hawk, Hawk).
  - If $v = c$, then there exists a unique Nash equilibrium, (Hawk, Hawk), though this is not a strict Nash equilibrium.
  - If $v < c$, then there exists three Nash equilibria: (Hawk, Dove) and (Dove, Hawk), which are non-symmetric strict equilibria, and a mixed strategy symmetric equilibrium.

Evolution in the Hawk-Dove Game

- If $v > c$, then we expect all agents to choose "Hawk". Those who do not will have lower fitness.
- A different way of thinking about the problem: imagine a population of agents playing "Dove" in this case.
- Suppose there is a **mutation**, so that one agent (or a small group of agents) starts playing "Hawk".
- This latter agent and its offspring will **invade** the population, because they will have greater fitness.
- The notion of **evolutionarily stable strategies** or **evolutionary stability** follows from this reasoning.

Evolutionarily Stable Strategies

- Let us first start with a general definition. For this, let us go back to mixed strategies in the strategic form game, except that this is a two player, symmetric game, so we write it simply as $\langle S, u \rangle$. A (possibly mixed) strategy is $\sigma \in \Sigma$.

Definition

(Evolutionarily stable strategy I) *A strategy $\sigma^* \in \Sigma$ is evolutionarily stable if there exists $\bar{\varepsilon} > 0$ such that for any $\sigma \neq \sigma^*$ (naturally with $\sigma \in \Sigma$) and for any $\varepsilon < \bar{\varepsilon}$, we have*

$$u(\sigma^*, \varepsilon\sigma + (1-\varepsilon)\sigma^*) > u(\sigma, \varepsilon\sigma + (1-\varepsilon)\sigma^*). \qquad \text{(Condition I)}$$

- Interpretation: strategy σ^* is evolutionarily stable if it cannot be invaded by any $\sigma \neq \sigma^*$. I.e., if, starting with a population playing σ^*, a small fraction $\varepsilon < \bar{\varepsilon}$ of agents play σ, then these players do worse (have lower fitness) than those playing σ^*.

Evolutionary Stability: Alternative Definition

- An alternative definition is:

Definition

(Evolutionarily stable strategy II) *A strategy* $\sigma^* \in \Sigma$ *is evolutionarily stable if for any* $\sigma \neq \sigma^*$ *(with* $\sigma \in \Sigma$*), we have*

1. $$u(\sigma^*, \sigma^*) \geq u(\sigma, \sigma^*).$$
2. *Moreover, if, for some* $\sigma \in \Sigma$, $u(\sigma^*, \sigma^*) = u(\sigma, \sigma^*)$, *then*
$$u(\sigma^*, \sigma) > u(\sigma, \sigma).$$

- Interpretation: An evolutionarily stable strategy σ^* is a Nash equilibrium. If σ^* is not a strict Nash equilibrium, then any other strategy σ that is a best response to σ^* must be worse against itself than against σ^*.

Evolutionary Stability: Equivalence of the Two Definitions

Theorem

The two definitions of evolutionarily stable strategies are equivalent.

Proof: (First implies second)

- Since the first definition holds for any $\varepsilon < \bar{\varepsilon}$, as $\varepsilon \to 0$, $u(\sigma^*, \varepsilon\sigma + (1-\varepsilon)\,\sigma^*) > u(\sigma, \varepsilon\sigma + (1-\varepsilon)\,\sigma^*)$ implies

$$u(\sigma^*, \sigma^*) \geq u(\sigma, \sigma^*),$$

thus establishing part 1 of the second definition.

Proof (continued)

- To establish part 2, suppose that $u(\sigma^*, \sigma^*) = u(\sigma, \sigma^*)$. Recall that u is linear in its arguments (since it is *expected* utility), so Condition I, $u(\sigma^*, \varepsilon\sigma + (1-\varepsilon)\,\sigma^*) > u(\sigma, \varepsilon\sigma + (1-\varepsilon)\,\sigma^*)$, can be written as

$$\varepsilon u(\sigma^*, \sigma) + (1-\varepsilon)\,u(\sigma^*, \sigma^*) > \varepsilon u(\sigma, \sigma) + (1-\varepsilon)\,u(\sigma, \sigma^*).$$

- Since $u(\sigma^*, \sigma^*) = u(\sigma, \sigma^*)$, this is equivalent to

$$\varepsilon u(\sigma^*, \sigma) > \varepsilon u(\sigma, \sigma),$$

Since $\varepsilon > 0$, part 2 of the second definition follows.

Proof (continued)

(Second implies first)

- We have that for any $\sigma \in \Sigma$, $u(\sigma^*, \sigma^*) \geq u(\sigma, \sigma^*)$. With the same argument as above, rewrite Condition I in the first definition as

$$\varepsilon u(\sigma^*, \sigma) + (1-\varepsilon)\, u(\sigma^*, \sigma^*) > \varepsilon u(\sigma, \sigma) + (1-\varepsilon)\, u(\sigma, \sigma^*). \qquad (*)$$

  If the inequality is strict, for ε sufficiently small, the first definition is satisfied (since $u(\sigma^*, \sigma) - u(\sigma, \sigma)$ is a finite number).

- If this relation holds as equality, then the second definition implies

$$u(\sigma^*, \sigma) > u(\sigma, \sigma).$$

  Multiply this by ε, use the fact that $u(\sigma^*, \sigma^*) = u(\sigma, \sigma^*)$, and add $(1-\varepsilon)\, u(\sigma^*, \sigma^*)$ to the left hand side and $(1-\varepsilon)\, u(\sigma, \sigma^*)$ to the right hand side, which gives $(*)$ and hence Condition I.

Evolutionary Stability and Nash Equilibrium

- Now given our second definition, the following is immediate:

Theorem

1. *A strict (symmetric) this Nash equilibrium of a symmetric game is an evolutionarily stable strategy.*
2. *An evolutionarily stable strategy is a Nash equilibrium.*

Proof: Both parts immediately follow from the second definition.

- Their converses are not true, however, as we will see.

Monomorphic and Polymorphic Evolutionarily Stability

- In addition, we could require an evolutionarily stable strategy (ESS) to be *monomorphic*—that is, all agents to use the same (pure) strategy.
- The alternative is *polymorphic*, where different strategies coexist, mimicking a mixed strategy equilibrium.
- With these definitions, let us return to the Hawk-Dove game.

The Hawk-Dove Game

Player 1/Player 2	Hawk	Dove
Hawk	$\left(\frac{1}{2}(v-c), \frac{1}{2}(v-c)\right)$	$(v, 0)$
Dove	$(0, v)$	$\left(\frac{1}{2}v, \frac{1}{2}v\right)$

- Recall that if $v > c$, then there is a unique **strict** Nash equilibrium, which is (Hawk, Hawk). Therefore, in this case "Hawk" is also an evolutionarily stable strategy.
- Moreover, it is monomorphic.

The Hawk-Dove Game (continued)

- What happens if $v = c$?
- Recall that now there is a unique Nash equilibrium, (Hawk, Hawk), which is not a strict Nash equilibrium.
- We will now show that it is still an evolutionarily stable strategy.
- Since "Dove" is also a best response, we need to look at $u(H, D)$ vs. $u(D, D)$. Clearly the first one is greater, so part 2 of the second definition is satisfied. Therefore "Hawk" is evolutionarily stable.
  - It is also monomorphic.
- This example also shows that strict Nash equilibrium is stronger than evolutionarily stable strategy.

The Hawk-Dove Game (continued)

- Suppose now $v < c$, then there exists three Nash equilibria: (Hawk, Dove) and (Dove, Hawk), which are non-symmetric strict equilibria, and a mixed strategy symmetric equilibrium.
- The first observation is that there exists no monomorphic evolutionarily stable strategy. This shows the importance of looking at polymorphic strategies.
- Since this is a random matching game, clearly the non-symmetric equilibria are irrelevant (why?).
- Is the mixed strategy equilibrium evolutionarily stable?

The Hawk-Dove Game (continued)

- First note that when $v < c$, the unique mixed strategy equilibrium of the strategic form game involves each player playing "Hawk" with probability v/c. The polymorphic evolutionary stable outcome will be a population where fraction v/c of the agents are type "Hawk". Let us designate this is by σ^*.
- We now need to show that such an outcome cannot be invaded by any other (mixed) strategy. That is, we need to check part 2 of the second definition. (Clearly, since σ^* is a mixed strategy, any other mixed strategy σ satisfies $u(\sigma^*, \sigma^*) = u(\sigma, \sigma^*)$).
- Consider a mixed strategy, $\sigma \neq \sigma^*$, where a fraction $p \neq v/c$ play "Hawk".

The Hawk-Dove Game (continued)

- Then

$$\begin{aligned} u(\sigma,\sigma) &= p^2 \times \frac{1}{2}(v-c) + p(1-p) \times v \\ &\quad + p(1-p) \times 0 + (1-p)^2 \times \frac{1}{2}v \\ u(\sigma^*,\sigma) &= \frac{v}{c}p \times \frac{1}{2}(v-c) + \frac{v}{c}(1-p) \times v \\ &\quad + \left(1-\frac{v}{c}\right)p \times 0 + \left(1-\frac{v}{c}\right)(1-p) \times \frac{1}{2}v \end{aligned}$$

- Therefore

$$u(\sigma^*,\sigma) - u(\sigma,\sigma) = \frac{1}{2}c\left(\frac{v}{c}-p\right)^2 > 0,$$

which establishes the desired result.

- This result also shows the possibility of polymorphic evolutionarily stable strategies.

Nash Equilibrium Does Not Imply ESS

- Consider the modified rock-paper-scissors game:

	R	P	S
R	(γ, γ)	$(-1, 1)$	$(1, -1)$
P	$(1, -1)$	(γ, γ)	$(-1, 1)$
S	$(-1, 1)$	$(1, -1)$	(γ, γ)

- Here $0 \leq \gamma < 1$. If $\gamma = 0$, this is the standard rock-paper-scissors game.
- For all such γ, there is a unique mixed strategy equilibrium $\sigma^* = (1/3, 1/3, 1/3)$, with expected payoff $u(\sigma^*, \sigma^*) = \gamma/3$. But for $\gamma > 0$, this is not ESS. For example, $\sigma = R$ would invade, since $u(\sigma, \sigma^*) = \gamma/3 < u(\sigma, \sigma) = \gamma$.
- This also shows that ESS doesn't necessarily exist.

Do Animals Play Games?

- The answer seems to be yes.
- They seem to play mixed strategies: sticklebacks are able to coordinate between the two sides of a fish tank, with different amounts of food "supply".
- This is like a "mixed strategy," since any food is shared among the sticklebacks at that end.
- Remarkably, when the relative amounts of food supplies into the fish tank at two sides is varied, sticklebacks are able to settle into the appropriate "mixed strategy" pattern given the new food supplies.

Do Animals Play Games?

- Side-blotched lizards seem to play a version of the Hawk-Dove game. Three productive strategies for male lizards with distinct throat colors (that are genetically determined):
  - orange color: very aggressive and defend large territories;
  - blue color: less aggressive defense smaller territories;
  - yellow color: not aggressive, opportunistic mating behavior.
- Tails seem to be as follows:
  - when all are orange, yellow does well; when all are yellow, blue does well; and when all are blue, orange does well.
- This is similar to the modified rock-paper-scissors pattern, and in nature, it seems that there are fluctuations in composition of male colorings as we should expect on the basis that the game does not have any evolutionarily stable strategies.
- See Karl Sigmund (1993) *Games of Life*, for many more fascinating examples, for animals, lower organisms and cells.

Dynamics

- The discussion of "dynamics" so far was largely heuristic.
- Are there actual dynamics of populations resulting from "game-theoretic" interactions that lead to evolutionarily stable strategies?
- Question at the intersection of game theory and population dynamics.
- The answer to this question is yes, and here we will discuss the simplest example, replicator dynamics.
- Throughout, we continue to focus on symmetric games.

Replicator Dynamics

- Let us formalize the discussion of fitness and offspring provided above.
- Let us enumerate the strategies by $s = 1, 2, ..., K$.
- Denote the fraction of the population playing strategy s by x_s.
- The setup is similar to that considered above: at each instant, each agent is randomly matched with another from a large population.
- What matters is expected fitness given by

$$u(s, \sigma).$$

In particular, recall that this is the expected fitness (payoff) of agents playing s when the mixed strategy induced by the polymorphic strategy profile is σ.

Replicator Dynamics (continued)

- Then, we can posit the following dynamic evolution:

$$x_s(t+\tau) - x_s(t) = x_s(t) \frac{\tau \left[u(s, \sigma(t)) - \bar{u}(\sigma(t))\right]}{\bar{u}(\sigma(t))},$$

(Replicator equation)

for each $s = 1, 2, ..., K$ and for all t and τ, where

$$\bar{u}(\sigma(t)) = \sum_{s=1}^{K} x_s(t)\, u(s, \sigma(t))$$

is average fitness at time t and $\sigma(t)$ is the vector of $x_s(t)$'s.

- Naturally, $\sum_{s=1}^{K} x_s(t) = 1$ by definition.
- This equation gives discrete time dynamics when $\tau = 1$. But the equation is valid for any τ.
- Intuitively, the greater is the fitness of a strategy relative to the average fitness, the greater is its relative increase in the population. Clearly, this equation is meaningful, i.e., $\sum_{s=1}^{K} x_s(t+\tau) = 1$.

Replicator Dynamics: Continuous Time

- It is most convenient to work with replicator dynamics in continuous time.
- Divide both sides of the replicator equation by τ and take the limit as $\tau \to 0$. This gives

$$\lim_{\tau \to 0} \frac{x_s(t+\tau) - x_s(t)}{\tau} = x_s(t) \frac{[u(s, \sigma(t)) - \bar{u}(\sigma(t))]}{\bar{u}(\sigma(t))}.$$

  Therefore

$$\dot{x}_s(t) = x_s(t) \frac{u(s, \sigma(t)) - \bar{u}(\sigma(t))}{\bar{u}(\sigma(t))}, \qquad \text{(Continuous replicator)}$$

  where recall that $\dot{x}_s(t) \equiv dx_s(t)/dt$.
- Notice that $x_s(t)$ is not written in the denominator of the left-hand side, since it can be equal to zero.

Replicator Dynamics: Implications

- Now we can think of the dynamics starting from an arbitrary distribution of strategies in the population.
- There are two ways of doing this:
  1. Ask whether a particular vector of distribution x^* is a **stationary state** of equation (Continuous replicator), meaning that it has $\dot{x}_s^*(t) = 0$ for all s.
  2. Ask whether a particular vector of distribution x^* is an **asymptotically stable state**, meaning that there exists a neighborhood of x^* such that starting from any x_0 in this neighborhood, dynamics induced by (Continuous replicator) approach x^*.

Replicator Dynamics and Nash Equilibria

Theorem

If x^* is a Nash equilibrium, then it is a stationary state.

Proof: If x^* is a Nash equilibrium, then it is a best response to itself, and thus no strategy has $u(s, \sigma(t)) - \bar{u}(\sigma(t)) > 0$, and $u(s, \sigma(t)) - \bar{u}(\sigma(t)) = 0$ only for strategies in the support of the mixed strategy profile induced by x^*. Thus for any s, either $u(s, \sigma(t)) - \bar{u}(\sigma(t)) = 0$ or $x_s(t) = 0$, and hence $\dot{x}_s^*(t) = 0$ for all s.

- However, the converse of this statement is not true, since if x^* corresponds to a non-Nash pure strategy, then $x_s^*(t) = 0$ for all s other than the pure strategy in question, and x^* is stable.
- Thus stability is not a particularly relevant concept. We would like x^* to be robust to "perturbations"—or against attempts at invasion. This requires *asymptotic stability.*

Replicator Dynamics and Nash Equilibria (continued)

Theorem

If x^* is asymptotically stable, then it is a Nash equilibrium.

- The proof is immediate if x^* corresponds to a pure strategy (monomorphic population).
- In the case where x^* corresponds to a mixed strategy Nash equilibrium, the proof is also straightforward but long. The basic idea is that equation (Continuous replicator) implies that we are moving in the direction of "better replies"—relative to the average. If this process converges, then there must not exist any more (any other) strict better replies, and thus we must be at a Nash equilibrium.
- The converse is again not true.

Replicator Dynamics and Nash Equilibria (continued)

- Consider, for example,

	A	B
A	$(1,1)$	$(0,0)$
B	$(0,0)$	$(0,0)$

- Here (B,B) is a Nash equilibrium, but clearly it is not asymptotically stable, since B is weakly dominated, and thus any perturbation away from (B,B) will start a process in which the fraction of agents playing A steadily increases.

Replicator Dynamics and Evolutionary Stability

- The key result here is the following, which justifies the focus on Evolutionary stable strategies.

Theorem

If x^* is evolutionarily stable, then it is asymptotically stable.

- The proof is again somewhat delicate, but intuitively straightforward. The first definition of ESS states that for small enough perturbations, the evolutionarily stable strategy is a strict best response. This essentially implies that in the neighborhood of the ESS σ^*, σ^* will do better than any other strategy σ, and thus according to (Continuous replicator), the fraction of those playing σ^* should increase, thus implying asymptotic stability.
- The converse of this result is not true, but mostly because of "technical reasons". Versions of its converse can be developed.

Evolution and Network Structure

- So far, no network structure in evolutionary interactions because of the **random matching** assumption (and this will be also the case when we turn to learning next).
- But this is not realistic. In practice, animals, organisms and humans play and compete more against "nearby" agents.
- One interesting area is to incorporate network structure into dynamics of game-theoretic behavior.
- We will do so in two different contexts later in the course:
  1. Evolution with local interactions.
  2. Games in which payoffs are determined by local interactions (so-called **Network Games**).

Learning vs. Evolution

- Evolution is a good model for fully myopic behavior. But even when individuals follow rules of thumb, they are not fully myopic.
- Moreover, in evolution, the time scales are long. We need "mutations," which are random and, almost by definition, rare.
- In most (human) game-theoretic situations, even if individuals are not fully rational, they can imitate more successful strategies quickly, and learn the behavior of their opponents and best respond to those.
- This suggests a related but distinct approach to dynamic game-theoretic behavior, which is taken in the literature on **learning in games**.
  - Note that this is different from Bayesian game-theoretic learning, which we will discuss later in the course.

Models of Learning

- One approach is to try to import menu of the insides of evolutionary game theory into the area of "learning in games".
- For example, for a symmetric game, we could posit **an imitation rule** that takes the form of equation (Continuous replicator), i.e., individuals imitate the strategies of others in proportion to how much they outperform the average in the population. Though plausible, this requires "global knowledge" on the part of individuals about others' payoffs.
- More importantly, in the context of learning, it may be more fruitful to ask: "what are players learning about?" The most plausible answer is the strategies of others.
- One of the earliest learning rules, fictitious play, introduced in Brown (1951) "Iterative solutions of games by fictitious play," is motivated by this type of reasoning.
- The idea is to look at a dynamic process where each player best responds to the time average of the behavior of its opponents.

Setup

- Let us first focus on a strategic form game $\langle \mathcal{I}, (S_i)_{i \in \mathcal{I}}, (u_i)_{i \in \mathcal{I}} \rangle$.
- The players play this game at times $t = 1, 2, \ldots$.
- The stage payoff of player i is again given by $u_i(s_i, s_{-i})$ (for pure strategy profile s_i, s_{-i}).
- For $t = 1, 2, \ldots$ and $i = 1, 2$, define the function $\eta_i^t : S_{-i} \to \mathbb{N}$, where $\eta_i^t(s_{-i})$ is the number of times player i has observed the action s_{-i} before time t. Let $\eta_i^0(s_{-i})$ represent a starting point (or fictitious past).
- For example, consider a two player game, with $S_2 = \{U, D\}$. If $\eta_1^0(U) = 3$ and $\eta_1^0(D) = 5$, and player 2 plays U, U, D in the first three periods, then $\eta_1^3(U) = 5$ and $\eta_1^3(D) = 6$.

The Basic Idea

- The be the idea of fictitious play is that each player assumes that his opponent is using a *stationary mixed strategy*, and updates his beliefs about this stationary mixed strategies at each step.
- Players choose actions in each period (or stage) to maximize that period's expected payoff given their prediction of the distribution of opponent's actions, which they form according to:

$$\mu_i^t(s_{-i}) = \frac{\eta_i^t(s_{-i})}{\sum_{\bar{s}_{-i} \in S_{-i}} \eta_i^t(\bar{s}_{-i})}.$$

- For example, in a two player game, player i forecasts player $-i$'s strategy at time t to be the empirical frequency distribution of past play.

Fictitious Play Model of Learning

- Given player i's belief/forecast about his opponents play, he chooses his action at time t to maximize his payoff, i.e.,

$$s_i^t \in \arg\max_{s_i \in S_i} u_i(s_i, \mu_i^t).$$

- Even though fictitious play is "belief based," it is also **myopic**, because players are trying to maximize current payoff without considering their future payoffs. Perhaps more importantly, they are also not learning the "true model" generating the empirical frequencies (that is, how their opponent is actually playing the game).

Example

- Consider the fictitious play of the following game:

	L	R
U	$(3,3)$	$(0,0)$
D	$(4,0)$	$(1,1)$

- Note that this game is dominant solvable (D is a strictly dominant strategy for the row player), and the unique NE (D, R).
- Assume that $\eta_1^0 = (3,0)$ and $\eta_2^0 = (1, 2.5)$. Then fictitious play proceeds as follows:
  - *Period 1:* Then, $\mu_1^0 = (1,0)$ and $\mu_2^0 = (1/3.5, 2.5/3.5)$, so play follows $s_1^0 = D$ and $s_2^0 = L$.
  - *Period 2:* We have $\eta_1^1 = (4,0)$ and $\eta_2^1 = (1, 3.5)$, so play follows $s_1^1 = D$ and $s_2^1 = R$.
  - *Period 3:* We have $\eta_1^1 = (4,1)$ and $\eta_2^1 = (1, 4.5)$, so play follows $s_1^2 = D$ and $s_2^2 = R$.
  - *Periods 4*:...

Example (continued)

- Since D is a dominant strategy for the row player, he always plays D, and μ_2^t converges to $(0, 1)$ with probability 1.
- Therefore, player 2 will end up playing R.
- The remarkable feature of the fictitious play is that players don't have to know anything about their opponent's payoff. They only form beliefs about how their opponents will play.

Convergence of Fictitious Play to Pure Strategies

- Let $\{\sigma^t\}$ be a sequence of strategy profiles generated by fictitious play (where for each t, $\sigma^t \in \Sigma^t$). Let us now study the asymptotic behavior of the sequence $\{\sigma^t\}$, i.e., the convergence properties of the sequence $\{\sigma^t\}$ as $t \to \infty$.
- We first define the notion of convergence to pure strategies.

Definition

The sequence $\{\sigma^t\}$ converges to s if there exists T such that $\sigma^t = s$ for all $t \geq T$ (i.e., it puts probability 1 on pure strategy s).

Theorem

Let $\{\sigma^t\}$ be a sequence of strategy profiles generated by fictitious play.

1. *If $\{\sigma^t\}$ converges to $\bar{s}$, then $\bar{s}$ is a pure strategy Nash equilibrium.*
2. *Suppose that for some t, $\sigma^t = s^*$, where s^* is a strict Nash equilibrium. Then $\sigma^\tau = s^*$ for all $\tau > t$.*

Proof

- Part part 1 is straightforward. Consider the proof of part 2.
- Let $\sigma^t = s^*$. We will show that $\sigma^{t+1} = s^*$. Note that

$$\mu_i^{t+1} = (1-\alpha)\mu_i^t + \alpha s_{-i}^t = (1-\alpha)\mu_i^t + \alpha s_{-i}^*,$$

  where, abusing the notation, we used s_{-i}^t to denote the degenerate probability distribution and

$$\alpha = \frac{1}{\sum_{s_{-i}} \eta_i^t(s_{-i}) + 1}.$$

- Therefore, by the linearity of the *expected utility*, we have for all $s_i \in S_i$,

$$u_i(s_i, \mu_i^{t+1}) = (1-\alpha)u_i(s_i, \mu_i^t) + \alpha u_i(s_i, s_{-i}^*).$$

- Since s_i^* maximizes both terms (in view of the fact that s^* is a strict Nash equilibrium), it follows that s_i^* will be played at $t+1$.

Convergence of Fictitious Play to Mixed Strategies

- The preceding notion of convergence only applies to pure strategies. We next provide an alternative notion of convergence, i.e., convergence of **empirical distributions or beliefs.**

Definition

The sequence $\{\sigma^t\}$ converges to $\sigma \in \Sigma$ in the time-average sense if for all i and for all $s_i \in S_i$, we have

$$\lim_{T \to \infty} \frac{[\text{number of times } s_i^t = s_i \text{ for } t \leq T]}{T+1} = \sigma(s_i),$$

i.e., $\mu_{-i}^T(s_i)$ converges to $\sigma(s_i)$ as $T \to \infty$.

Convergence in Matching Pennies: An Example

Player 1 \ Player 2	heads	tails
heads	$(1,-1)$	$(-1,1)$
tails	$(-1,1)$	$(1,-1)$

Time	η_1^t	η_2^t	Play
0	$(0,0)$	$(0,2)$	(H,H)
1	$(1,0)$	$(1,2)$	(H,H)
2	$(2,0)$	$(2,2)$	(H,T)
3	$(2,1)$	$(3,2)$	(H,T)
4	$(2,2)$	$(4,2)$	(T,T)
5	$(2,3)$	$(4,3)$	(T,T)
6	...	...	(T,H)

- In this example, play continues as a deterministic cycle. The time average converges to the unique Nash equilibrium, $\big((1/2,1/2),(1/2,1/2)\big)$.

More General Convergence Result

Theorem

Suppose a fictitious play sequence $\{s^t\}$ converges to σ in the time-average sense. Then σ is a Nash equilibrium.

Proof:

- Suppose s^t converges to σ in the time-average sense.
- Suppose, to obtain a contradiction, that σ is not a Nash equilibrium. Then there exist some i, s_i, $s_i' \in S_i$ with $\sigma_i(s_i) > 0$ such that

$$u_i(s_i', \sigma_{-i}) > u_i(s_i, \sigma_{-i}).$$

Proof (continued)

- Choose $\varepsilon > 0$ such that

$$\varepsilon < \frac{1}{2}\Big[u_i(s_i', \sigma_{-i}) - u_i(s_i, \sigma_{-i})\Big],$$

and T sufficiently large that for all $t \geq T$, we have

$$\Big|\mu_i^T(s_{-i}) - \sigma_{-i}(s_{-i})\Big| < \frac{\varepsilon}{\max_{s \in S} u_i(s)},$$

which is possible since $\mu_i^t \to \sigma_{-i}$ by assumption.

Proof (continued)

- Then, for any $t \geq T$, we have

$$\begin{aligned} u_i(s_i, \mu_i^t) &= \sum_{s_{-i}} u_i(s_i, s_{-i}) \mu_i^t(s_{-i}) \\ &\leq \sum_{s_{-i}} u_i(s_i, s_{-i}) \sigma_{-i}(s_{-i}) + \varepsilon \\ &< \sum_{s_{-i}} u_i(s_i', s_{-i}) \sigma_{-i}(s_{-i}) - \varepsilon \\ &\leq \sum_{s_{-i}} u_i(s_i', s_{-i}) \mu_i^t(s_{-i}) = u_i(s_i', \mu_i^t). \end{aligned}$$

- This shows that after T, s_i is never played, implying that as $T \to \infty$, $\mu_{-i}^t(s_i) \to 0$. But this contradicts the fact that $\sigma_i(s_i) > 0$, completing the proof.

Convergence

Theorem

Fictitious play converges in the time-average sense for the game G under any of the following conditions:

- *G is a two player zero-sum game.*
- *G is a two player nonzero-sum game where each player has at most two strategies.*
- *G is solvable by iterated strict dominance.*
- *G is an identical interest game, i.e., all players have the same payoff function.*
- *G is a potential game.*

- Below, we will prove convergence (in a stronger sense than here) in potential games using continuous-time fictitious play.

Miscoordination

- However, convergence in the time-average sense is not necessarily a natural convergence notion, as illustrated in the following example.
- Consider the fictitious play of the following game:

Player 1 \ Player 2	A	B
A	$(1,1)$	$(0,0)$
B	$(0,0)$	$(1,1)$

- Note that this game had a unique mixed Nash equilibrium $\big((1/2,1/2),(1/2,1/2)\big)$.

Miscoordination (continued)

- Consider the following sequence of play:

Time	η_1^t	η_2^t	Play
0	$(1/2, 0)$	$(0, 1/2)$	(A, B)
1	$(1/2, 1)$	$(1, 1/2)$	(B, A)
2	$(3/2, 1)$	$(1, 3/2)$	(A, B)
3	...	...	(B, A)
4	...	...	(A, B)

- Play continues as (A,B), (B,A), . . ., which is again a deterministic cycle. The time average converges to $\big((1/2, 1/2), (1/2, 1/2)\big)$, which is a mixed strategy equilibrium of the game. But players never successfully coordinate and receive zero payoffs throughout!

Non-convergence

- Convergence of fictitious play can also not be guaranteed.
- Shapley showed that in modified rock-scissors-paper game, fictitious play does not converge.
- Recall:

	R	P	S
R	(γ, γ)	$(-1, 1)$	$(1, -1)$
P	$(1, -1)$	(γ, γ)	$(-1, 1)$
S	$(-1, 1)$	$(1, -1)$	(γ, γ)

- When $\gamma = 0$ this is a zero-sum game and there is convergence to a deterministic cycle as in the matching pennies. When $\gamma > 0$, Shapley showed that there are "cycles" of ever-increasing length, thus non-convergence.

Continuous-Time Fictitious Play

- As with the replicator dynamics, continues-time version of fictitious play is more tractable.
- To show that in potential games, fictitious play converges factual Nash equilibrium behavior, we will use continuous-time fictitious play.
- Denote the empirical distribution of player i's play up to (but not including) time t when time intervals are of length Δt by

$$p_i^t(s_i) = \frac{\sum_{\tau=0}^{(t-\Delta t)/\Delta t} \mathcal{I}\{s_i^\tau = s_i\}}{t}.$$

- We use $p^t \in \Sigma$ to denote the product distribution formed by the p_i^t.
- We can now think of making time intervals Δt smaller as we did in replicator dynamics (also rescaling time), which will lead us to a version a fictitious play in continuous time. We next study this continuous-time fictitious play model.

Continuous-Time Fictitious Play (continued)

- In continuous time fictitious play (CTFP), the empirical distributions of the players are updated in the direction of a best response to their opponents' past action:

$$\frac{dp_i^t}{dt} \in BR_i(p_{-i}^t) - p_i^t,$$

where

$$BR_i(p_{-i}^t) = \arg\max_{\sigma_i \in \Sigma_i} u_i(\sigma_i, p_{-i}^t).$$

- In addition, we impose that

$$\frac{dp_i^t}{dt} = 0 \text{ if } p^t \text{ is a Nash equilibrium.}$$

- We next show that fictitious play converges for (finite) potential games.

Convergence of Fictitious Play for Potential Games

- Recall that a function $\Phi : S \to \mathbb{R}$ is an **exact potential function** for the game G if for each $i \in \mathcal{I}$ and all $s_{-i} \in S_{-i}$,

$$u_i(x, s_{-i}) - u_i(z, s_{-i}) = \Phi(x, s_{-i}) - \Phi(z, s_{-i}), \text{ for all } x, z \in S_i.$$

- Here we focus on exact potential games, but the result is straightforward to generalize to ordinal potential games.
- Consider the continuous time fictitious play (CTFP) dynamics:

$$\frac{dp_i^t}{dt} \in BR_i(p_{-i}^t) - p_i^t.$$

- Let $\{p_i^t\}$ denote the sequence generated by CTFP dynamics and let $\sigma_i^t = p_i^t + dp_i^t/dt$. Note that $\sigma_i^t \in BR_i(p_{-i}^t)$.

Theorem

In finite potential games, continuous-time fictitious play converges to equilibrium behavior.

Proof

- For each player i, we define the function

$$U_i(\sigma_i, \sigma_{-i}) = \max_{\sigma_i' \in \Sigma} u_i(\sigma_i', \sigma_{-i}) - u_i(\sigma_i, \sigma_{-i}),$$

- Intuitively, the function U_i gives the maximum possible payoff improvement player i can achieve by a unilateral deviation.
- We define $W(t) \equiv \sum_i U_i(p^t)$. Observe that

$$\begin{aligned}
\frac{d}{dt}(\Phi(p^t)) &= \frac{d}{dt}\left[\sum_{s_i \in S_i} \cdots \sum_{s_n \in S_n} p_1^t(s_1) \cdots p_n^t(s_n)\Phi(s)\right] \\
&= \sum_i \sum_{s_i \in S_i} \cdots \sum_{s_n \in S_n} \frac{dp_i^t}{dt}(s_i)\left(\prod_{j \neq i} p_j^t(s_j)\right)\Phi(s) \\
&= \sum_i \Phi\left(\frac{dp_i^t}{dt}, p_{-i}^t\right).
\end{aligned}$$

Proof (Continued)

- The preceding explicit derivation essentially follows from the fact that Φ is linear in its arguments, because these are mixed strategies of players. Therefore, the time derivative can be directly applied to the arguments.
- Now, observe that

$$\Phi\left(\frac{dp_i^t}{dt}, p_{-i}^t\right) = \Phi(\sigma_i^t - p_i^t, p_{-i}^t) = \Phi(\sigma_i^t, p_{-i}^t) - \Phi(p^t) = U_i(p^t),$$

where the second equality again follows by the linearity of Φ in mixed strategies. The last equality uses the fact that $\sigma_i^t \in BR_i(p_{-i}^t)$.
- Combining this relation with the previous one, we have

$$\frac{d}{dt}(\Phi(p^t)) = \sum_i U_i(p^t) = W(t).$$

Proof (Continued)

- Since $W(t)$ is nonnegative everywhere, we conclude $\Phi(p^t)$ is nondecreasing as t increases; thus $\Phi^* = \lim_{t\to\infty} \Phi(p^t)$ exists (since Φ is bounded above, $\Phi^* < \infty$).
- Moreover, we have

$$\Phi^* - \Phi(p^t) \geq \Phi(p^{t+\Delta}) - \Phi(p^t) = \int_0^{\Delta} W(t+\tau)d\tau \geq 0.$$

  - the first inequality uses the fact that since Φ is nondecreasing; the middle inequality follows from the fundamental theorem of calculus, and the last inequality simply uses the fact that $W(t)$ is everywhere nonnegative.
- Since the left-hand side converges to zero, we conclude that $W(t) \to 0$ as $t \to \infty$.
- This establishes that for each i and for any initial condition p^0,

$$\lim_{t\to\infty} \left[\max_{\sigma_i' \in \Sigma_i} \Phi(\sigma_i', p_{-i}^t) - \Phi(p_i^t, p_{-i}^t) \right] = 0.$$

Proof (Continued)

- Since Φ is the potential function, this implies

$$\lim_{t\to\infty}\left[\max_{\sigma_i'\in\Sigma_i} u_i(\sigma_i', p_{-i}^t) - u_i(p_i^t, p_{-i}^t)\right] = 0.$$

- Therefore, behavior converges to the equilibrium.
- Notice that what we have here is much stronger than convergence of fictitious play in empirical distribution (the results discussed above).
- Instead, we have that for any initial condition p^0, p^t converges to a set of empirical distributions P^∞, where $\Phi(p) = \Phi^*$ for all $p \in P^\infty$, and the mixed strategy of each player is the one that maximizes payoff in response to these distributions.
- **Implication:** the miscoordination illustrated above cannot happen.
- Moreover, recall that potential games have pure strategy equilibria. If this pure strategy equilibrium is the unique equilibrium, this result implies convergence to this unique equilibrium.

Implications

- This result implies that in potential games, rule of thumb behavior will take us towards Nash equilibrium.
- While this result is stated for finite games, it can be generalized for infinite games as well.
- Since, as we have seen, many congestion, network traffic and routing, and network formation games are potential games, these results imply that for a range of network games, Nash equilibrium behavior will emerge even without very sophisticated reasoning on the part of the players.