Learning

the traffic game

focus on longer term learning and implications for “steady states”

active vs. passive learning

for active learning players must be patient so willing to undertake investment

The Mystery in Human Learning

- not why people learn so badly – why they learn so well.
- behavioral economists, psychologists, economists and computer scientists model human learning using naïve and primitive models.
- models designed by computer scientists to make the best possible decisions cannot come close to the learning ability of the average human child, chimpanzee or even rat.
- equilibrium models and rational expectations: if we have to choose between best models of learning and perfect learning – for most situations of interest to economists perfect learning fits the facts better

Global Convergence?

- “grail” of learning research: global convergence theorem for convincing learning processes
- easy to construct examples of learning processes that don’t converge
- non-convergence looks like cob-web; people repeat the same mistakes over and over; not terrifically plausible
- we seem to see much “equilibriumness” around us: traffic, refugee camps

Overview

- stochastic procedures can be globally stable: fishing for Nash equilibrium
- here properties of basic passive learning procedures

Worst-case or Universal analysis vs. Bayesian analysis

- opponents may be smarter than you
- their process of optimization may result in play not in the support of your prior
- probability 1 with respect to your own beliefs is not meaningful in the setting of a game
- example: everyone believing that they face a stationary process (a common statistical assumption) implies that no one will actually behave in a stationary way
- these deficiencies in the robustness of Bayes learning are why there is no satisfactory global convergence theorem for learning procedures

"Classical" Case of Fictitious Play

- keep track of frequencies of opponents' play
- begin with an initial or prior sample
- play a best-response to historical frequencies
- not well defined if there are ties, but for generic payoff/prior there will be no ties
- optimal procedure against i.i.d. opponents
- how well does fictitious play do if the i.i.d. assumption is wrong?

How well can fictitious play do in the long-run?

- notice that fictitious play only keeps track of frequencies: can fictitious play do as well in the long-run as if those frequencies (but not the order of the sample) was known in advance?
- alternatively: suppose that a player is constrained to play the same action in every period, so that he does not care about the order of observations

Universal Consistency

let u_t^i be actual utility at time *t*

let ϕ_t^{-i} be frequency of opponents' play (joint distribution over S^{-i})

suppose that for *all* (note that this does not say "for almost all") sequences of opponent play

$$\liminf_{T\to\infty}(1/T)\sum_{t=1}^{T}u_t^i - \max_{s^i} u^i(s^i,\phi_T^{-i}) \geq 0$$

then the learning procedure is universally consistent

Is fictitious play universally consistent? Fudenberg and Kreps example

0,0	1,1
1,1	0,0

this coordination game is played by two identical players

suppose they use *identical deterministic* learning procedures

then they play UL or DR and get 0 in every period

this is not individually rational, let alone universally consistent

Theorem [Monderer, Samet, Sela; Fudenberg, Levine]: fictitious play is consistent provided the frequency with which the player switches strategies goes to zero

Self Confirming Equilibrium

$s_i \in S_i$ pure strategies for *i;* $\sigma_i \in \Sigma_i$ mixed

H_i information sets for *i*

$\overline{H}(\sigma)$ reached with positive probability under σ

$\pi_i \in \Pi_i$ behavior strategies

$\hat{\pi}(h_i \,|\, \sigma_i)$ map from mixed to behavior strategies

$\hat{\rho}(\pi)$, $\hat{\rho}(\sigma) \equiv \hat{\rho}(\hat{\pi}(\sigma))$ distribution over terminal nodes

μ_i a probability measure on Π_{-i}

$u_i(s_i \,|\, \mu_i)$ preferences

$$\Pi_{-i}(\sigma_{-i} \,\big|\, J) \equiv \{\pi_{-i} \,\big|\, \pi_i(h_i) = \hat{\pi}(h_i \,|\, \sigma_i), \forall h_i \in H_{-i} \cap J\}$$

Notions of Equilibrium

Nash equilibrium

a mixed profile σ such that for each $s_i \in \mathrm{supp}(\sigma_i)$ there exist beliefs μ_i such that

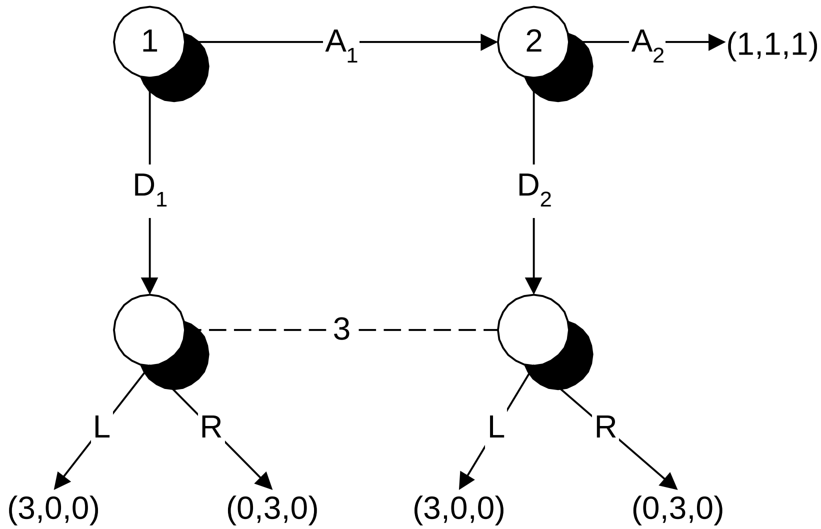
- s_i maximizes $u_i(\cdot|\mu_i)$
- $\mu_i(\Pi_{-i}(\sigma_{-i}|H)) = 1$

Unitary Self-Confirming Equilibrium

- $\mu_i(\Pi_{-i}(\sigma_{-i} \mid \bar{H}(\sigma))) = 1$

(=Nash with two players)

Fudenberg-Kreps Example



A_1, A_2 is self-confirming, but not Nash

any strategy for 3 makes it optimal for either 1 or 2 to play down

but in self-confirming, 1 can believe 3 plays R; 2 that he plays L

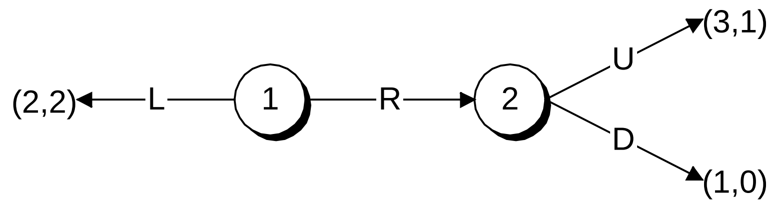
Heterogeneous Self-Confirming equilibrium

- $\mu_i(\Pi_{-i}(\sigma_{-i} \mid \bar{H}(s_i, \sigma))) = 1$

Can summarize by means of "observation function"

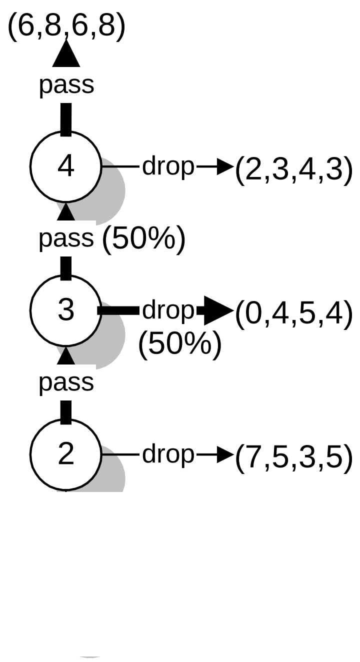
$$J(s_i, \sigma) = H, \bar{H}(\sigma), \bar{H}(s_i, \sigma)$$

Public Randomization



Remark: In games with perfect information, the set of heterogeneous self-confirming equilibrium payoffs (and the probability distributions over outcomes) are convex

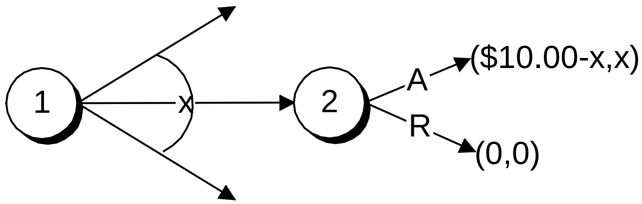
Example Without Public Randomization



Ultimatum Bargaining

player 1 proposes how to divide $10 in nickles

player 2 may accept or reject



Nash: any proposal by player 1 with all poorer proposals rejected and equal or better proposals accepted

Subgame Perfect: First player gets at least $9.95

US Data for Ultimatum

x	*Offers*	*Rejection Probability*
$2.60	3	33%
$4.25	13	18%
$5.00	13	0%
	29	

US $10.00 stake games, round 10

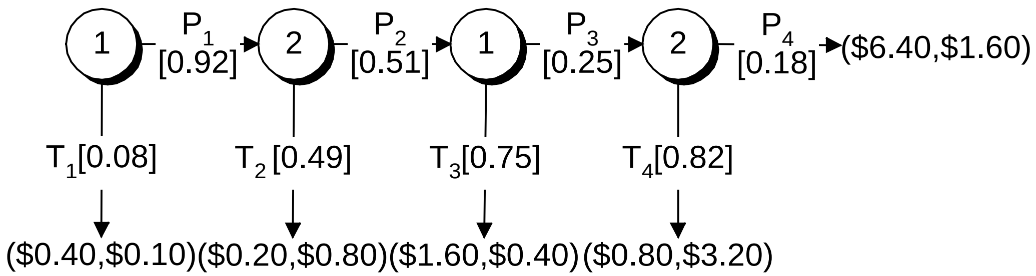
Trials	Rnd	Cntry Stake	Case	Expected Loss			Max Gain	Ratio
				Pl 1	Pl 2	Both		
27	10	US	H	$0.00	$0.67	$0.34	$10.00	3.4%
27	10	US	U	$1.30	$0.67	$0.99	$10.00	9.9%
10	10	USx3	H	$0.00	$1.28	$0.64	$30.00	2.1%
10	10	USx3	U	$6.45	$1.28	$3.86	$30.00	12.9%
30	10	Yugo	H	$0.00	$0.99	$0.50	$10?	5.0%
30	10	Yugo	U	$1.57	$0.99	$1.28	$10?	12.8%
29	10	Jpn	H	$0.00	$0.53	$0.27	$10?	2.7%
29	10	Jpn	U	$1.85	$0.53	$1.19	$10?	11.9%
30	10	Isrl	H	$0.00	$0.38	$0.19	$10?	1.9%
30	10	Isrl	U	$3.16	$0.38	$1.77	$10?	17.7%
	WC		H			$5.00	$10.00	50.0%

Rnds=Rounds, WC=Worst Case, H=Heterogeneous, U=Unitary

Comments on Ultimatum

- every offer by player 1 is a best response to beliefs that all other offers will be rejected so player 1's heterogeneous losses are always zero.
- big player 1 losses in the unitary case
- player 2 losses all knowing losses from rejected offers; magnitudes indicate that subgame perfection does quite badly
- as in centipede, tripling the stakes increases the size of losses a bit less than proportionally (losses roughly double).

Centipede Game: Palfrey and McKelvey



Numbers in square brackets correspond to the observed conditional probabilities of play corresponding to rounds 6-10, stakes 1x below.

This game has a unique self-confirming equilibrium; in it player 1 with probability 1 plays T_1

Trials / Rnd	Rnds	Stake	Case	Expected Loss			Max Gain	Ratio
				Pl 1	Pl 2	Both		
29*	6-10	1x	H	$0.00	$0.03	$0.02	$4.00	0.4%
29*	6-10	1x	U	$0.26	$0.17	$0.22	$4.00	5.4%
	WC	1x	H			$0.80	$4.00	20.0%
29	1-10	1x	H	$0.00	$0.08	$0.04	$4.00	1.0%
10	1-10	4x	H	$0.00	$0.28	$0.14	$16.00	0.9%

Rnds=Rounds, WC=Worst Case, H=Heterogeneous, U=Unitary

Comments on Experimental Results

- heterogeneous loss per player is small; because payoffs are doubling in each stage, equilibrium is very sensitive to a small number of player 2's giving money away at the end of the game.
- unknowing losses far greater than knowing losses
- quadrupling the stakes very nearly causes $\bar{\varepsilon}$ to quadruple
- theory has substantial predictive power: see WC
- losses conditional on reaching the final stage are quite large--inconsistent with subgame perfection. McKelvey and Palfrey estimated an incomplete information model where some "types" of player 2 liked to pass in the final stage. This cannot explain many players dropping out early so their estimated model fits poorly.

Self-confirming Equilibrium and Economic Policy

- example adapted from Sargent, Williams and Zhao [2006a] by Fudenberg and Levine [2009]
- game between a government and typical or representative consumer
- first: government chooses high or low inflation
- second: consumers choose high or low unemployment
- consumers always prefer low unemployment
- government prefers low inflation to high inflation, but cares more about unemployment being low than about inflation
- “full” rationality (subgame perfection): the consumer will always choose low unemployment; government recognizing this will always choose low inflation

Self-confirming Equilibrium

- government believes incorrectly that low inflation leads to high unemployment
- widespread belief at one time
- care more about employment than inflation so keep inflation high
- never learn that beliefs about low inflation are false
- in practice information about consequences of low inflation generated by policy “mistakes” and random shocks
- Sargent, Williams and Zhao [2006a] use sophisticated dynamic model of learning to analyze how U.S. Federal Reserve policy evolved post World War II to ultimately result in the conquest of U.S. inflation
- In particular they explain why it took so long – a cautionary note for economic policy makers

Keynes Beauty Contest

His explanation of how stock markets work

- players must choose the most beautiful woman from six photographs
- players who pick the most popular face win

a boring coordination game: every face is a Nash equilibrium

Keynes made a fortune for King's College Cambridge through his stock market investments; also lost it periodically, so he was lucky to quit while ahead

Nagel Experiment

choose a number between 0 and 100.

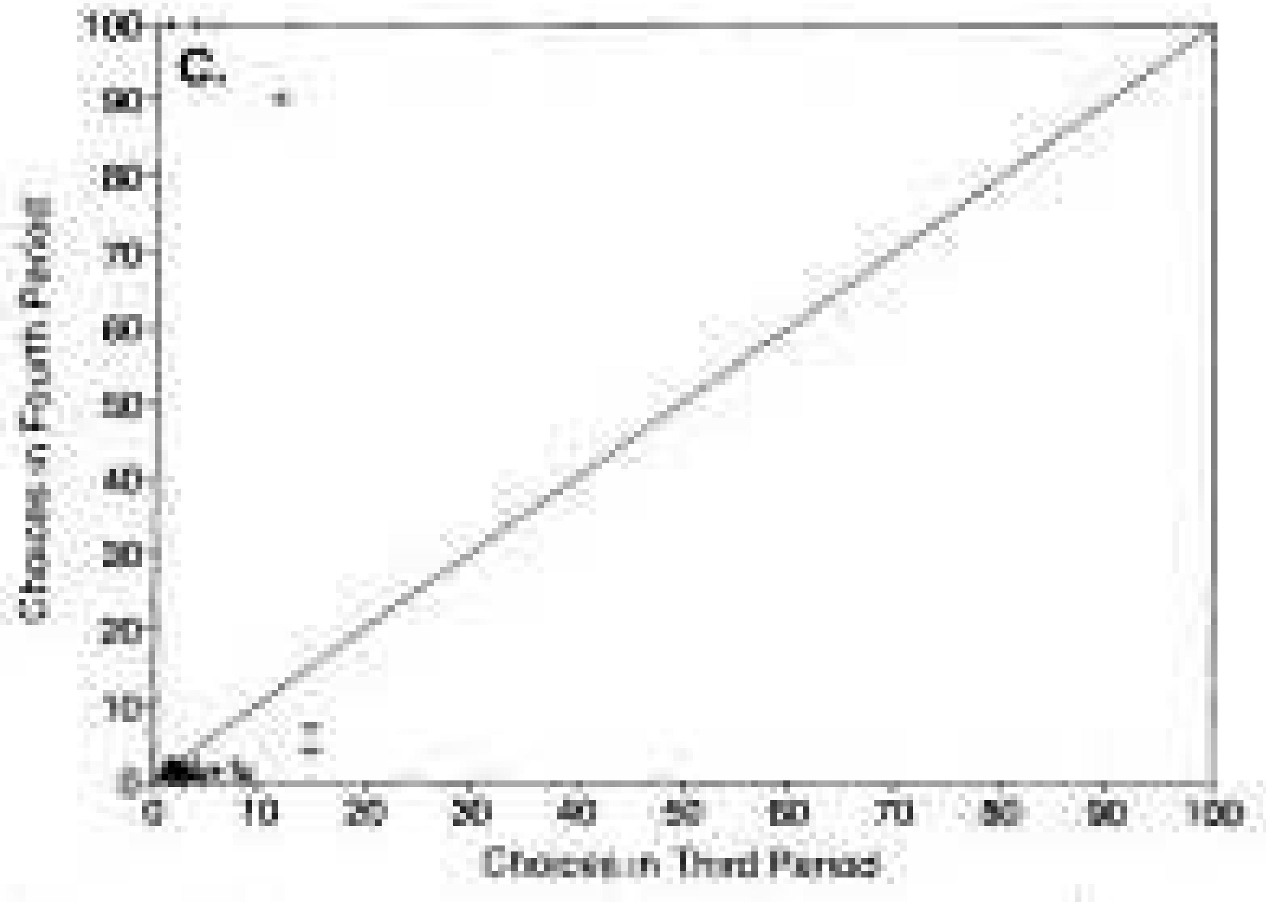
players closest to half the average value win

what you want to do depends on what you think average opinion is

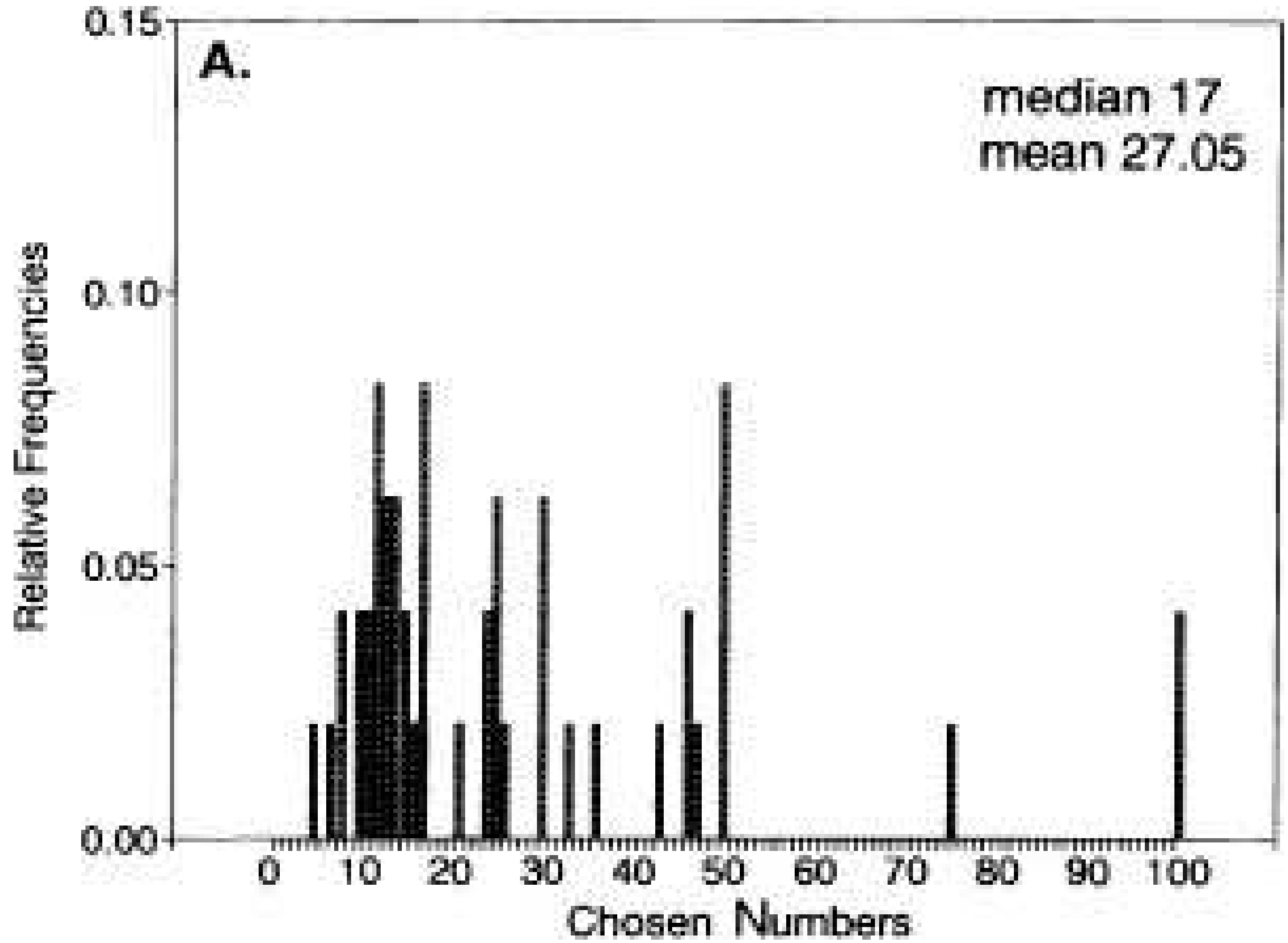
if you think that people choose randomly, average should be 50, so guess 25

(almost) a unique Nash equilibrium: zero and one

Third and Fourth Plays



First Time Play



Level-k Theory

One time play: not Nash equilibrium

Not expected to be Nash since one-time play

another theory does do a good job of explaining what is going on

originates with Nagel, developed by Stahl and Wilson [1994], Costa-Gomes., Crawford. and Broseta [2001] and Camerer, Ho and Chong [2004] and others

people differ in their sophistication

- naïve individuals – level-0 – play randomly
- less naïve individuals – level-1 – believe that their opponents are of level-0
- in general higher level and more sophisticated individuals – level-*k* – believe that they face a mixture of less sophisticated individuals – people with lower levels of *k*

How Long Does it Take to Learn?

The Hijacking Game (from Wikipedia)

Before September 11, 2001 pilots and flight attendants were trained in the FAA approved "Common Strategy"

- comply with hijackers' demands, land safely, let security forces handle situation
- advise passengers to sit quietly in order to increase chances of survival
- do not be a hero and endanger the passengers

the longer a hijacking persisted of a peaceful ending

1988-1997 an average of 18 hijackings per year, the vast majority ending peacefully

well-established, successful procedure, well validated by experience

The Rule Change:September 11, 2001

8:46 am: American Airlines Flight 11 crashes into the North Tower of the World Trade Center

9:03 am: United Flight 175 crashes into the South Tower of the World Trade Center

9:28 am: United Airlines Flight 83 is hijacked

9:37 am: American Airlines Flight 77 crashes into the West side of the Pentagon

9:57 am: passengers on United Airlines Flight 93 assault hijackers

So: an hour and 11 minutes

Since that time passengers have resisted essentially all hijackings

Global Convergence Once Again

(Foster and Young)

An individual is *stable* or *unstable*

If everyone played the same last period as this period, and you played a best response, there is a probability $0 < \mu < 1$ that you become stable

If you are stable, you play the same as last period

If you are unstable you play every choice with probability at least $\varepsilon > 0$

Suppose there is a pure strategy Nash equilibrium

- a stable Nash equilibrium is absorbing
- if we are not at Nash somebody is unstable
- if somebody is unstable there is a positive probability everyone will be unstable next period
- if everyone is unstable there is a positive probability there will be a Nash equilibrium played next period
- if a Nash equilibrium is played there is a positive probability everyone becomes stable
- hence all non-Nash and all not everybody stable are transitory
- we converge with probability one to a stable pure strategy Nash equilibrium
- how to deal with mixed strategy equilibria? Must have explicit mixing