Calibrated Learning and Global Convergence

David K. Levine

September 29, 2001

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 example)
- "full Bayes learning" (Kalai-Lehrer) results in convergence to Nash equilibrium
- Peyton just argued that such learning isn't really possible
- I'll try to convince you that "all sensible" learning procedures lead in the long-run to correlated equilibrium

- I'll start by motivating learning processes from an individual perspective (i.e. processes that "work")
- I'm only going to talk about pure forecasting (no causality)

"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? Notice the weakening of the criterion

- 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

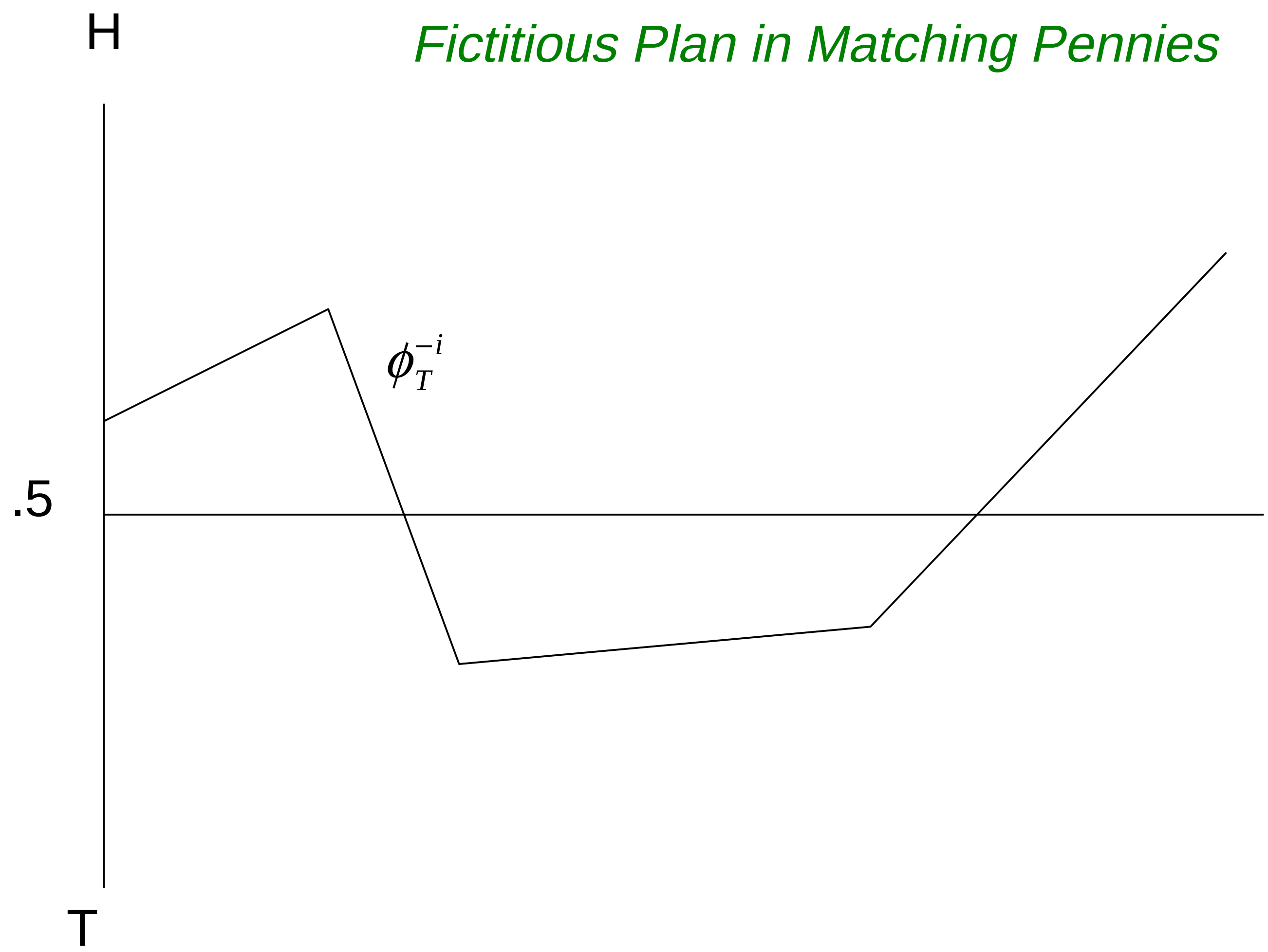
this coordination game is played by two identical players

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

Fictitious Plan in Matching Pennies



Smooth Fictitious Play

instead of maximizing $u^i(s^i, \phi^i_{t-1})$ maximize $u^i(\sigma^i, \phi^i_{t-1}) + \lambda v^i(\sigma^i)$

where v^i is smooth, concave and has derivatives that are unbounded at the boundary of the unit simplex

example: the *entropy* $v^i(\sigma^i) = -\sum_{s^i} \sigma^i(s^i) \log \sigma^i(s^i)$

as $\lambda \to 0$ this results in an approximate optimum to the original problem

however the solution to $u^i(\sigma^i, \phi^i_{t-1}) + \lambda v^i(\sigma^i)$ is smooth and interior (always puts positive weight on all pure strategies)

Theorem [Blackwell, Hannan, Fudenberg and Levine and others]: smooth fictitious play is ε universally consistent with $\varepsilon \to 0$ as $\lambda \to 0$

Calibration

Notice that pattern recognition is ruled out

Instead, use conditional probabilities; specifically

$$\phi_T^{-i}(\tilde{s}^i)$$

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

Interpetation of Calibration

weather forecasting example: calibrated beliefs, versus calibrated actions

consequence of universal calibration: global convergence to the set of correlated equilibria

Foster and Vohra: there are universally calibrated algorithms

How to do it?

$\hat{\sigma}^i(\phi)$ smooth fictitious play

suppose you play $\tilde{\sigma}^i$

with probability $\tilde{\sigma}^i(s^i)$ you play s^i

if you choose s^i then you "should" play $\hat{\sigma}^i(\phi_{t-1}^{-i}(s^i))$

so overall, you "should" play $\sum_{s^i} \tilde{\sigma}^i(s^i)\hat{\sigma}^i(\phi_{t-1}^{-i}(s^i))$

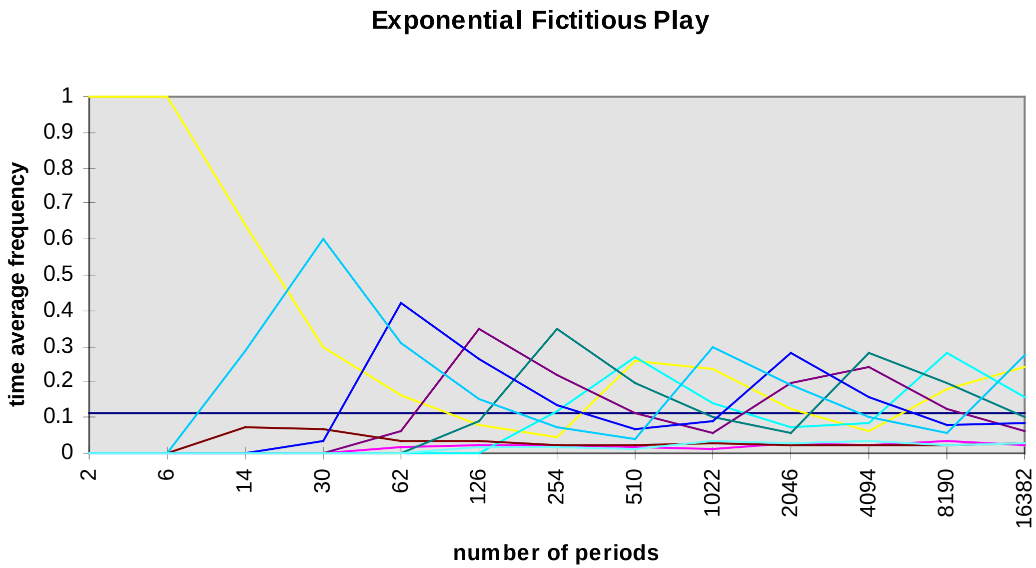
a fixed point problem then: $\tilde{\sigma}^i(s^i) = \sum_{s^i} \tilde{\sigma}^i(s^i)\hat{\sigma}^i(\phi_{t-1}^{-i}(s^i))$

easy to solve, and indeed the solution is calibrated

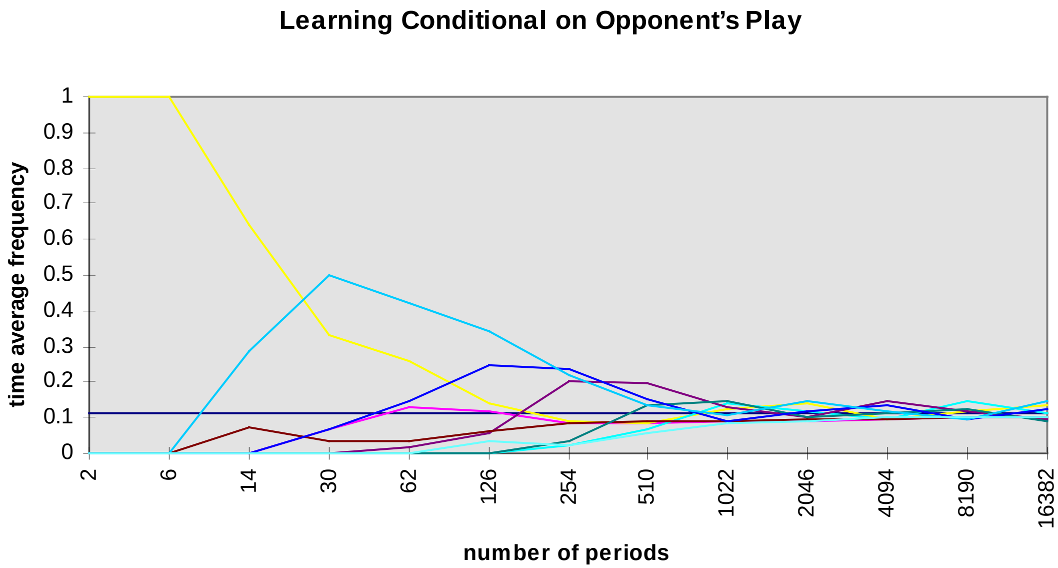
Shapley Example

	A	M	B
A	0,0	0,1	1,0
M	1,0	0,0	0,1
B	0,1	1,0	0,0

smooth fictitious play (time in logs)



condition on opponents last period play (time in logs)



Learning Conditional on Opponent's Play
time average frequency
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
6
14
30
62
126
254
510
1022
2046
4094
8190
16382
number of periods