Learning in Games

Steady State Learning

David K. Levine

October 5, 2010

Overview of the Model

- Capture ongoing learning in a steady state framework
- How? Knowledge is “lost”

- society consists of overlapping generations of finitely lived players
- indoctrinated into the social norm as children
- enter the world as young adults with prior beliefs that the social norm is true
- being young and relatively patient, having some residual doubt about the truth of what they were taught, and being rational Bayesians, young players optimally experiment to see what will happen

Rational Steady-State Learning

The Agent's Decision Problem

"agent" in the role of player *i* expects to play game T times wishes to maximize

$$\frac{1-\delta}{1-\delta^T} E \sum_{t=1}^{T} \delta^{t-1} u_t$$

u_t realized stage game payoff

agent believes that he faces a fixed time invariant probability distribution of opponents' strategies, unsure what the true distribution is

Definition 5.1: *Beliefs* μ_i *are* non-doctrinaire *if* μ_i *is given by a continuous density function* g_i *strictly positive at interior points.*

Note that allow priors can go to zero on the boundary, as is the case for many Dirichlet priors

The Policy Functiion

assume non-doctrinaire prior g_i^0

$g_i(\cdot \mid z)$ posterior starting with prior g_i after z is observed

agents are assumed to play optimally

(can write down dynamic programming problem)

histories are Y_i

optimal policy a map $r_i : Y_i \to S_i$ (may be several)

Steady States in an Overlapping generations model

- ♦ a continuum population
- ♦ doubly infinite sequence of periods
- ♦ generations overlap
- ♦ $1/T$ players in each generation
- ♦ $1/T$ enter to replace the $1/T$ player who leave
- ♦ each agent is randomly and independently matched with one agent from each of the other populations

each population assumed to use a common optimal rule r_i

look for a population steady state in which the fractions of each population playing pure strategies is time invariant

The Meeting Function

state at t is fraction of the population with each possible history $\theta_t^i(y_\tau^i)$.
for any state θ define number of people playing s^i:

$$\bar{\theta}^i(s^i) \equiv \sum_{y_\tau^i | r^i(y_\tau^i) = s^i} \theta^i(y_\tau^i)$$

$f^i(\theta)[y_\tau^i]$ fraction of population i with private history y_τ^i time $t+1$ when state at t was θ; that is, you meet people according to θ, what experience will result?

Details

of existing population $\theta^i(y^i_\tau)$ with particular history, fraction having experience $(y^i_\tau, r^i(y^i_\tau), z)$ is fraction that were matched with opponents playing strategies that led to the terminal node z

$\hat{s}^{-i}(s^i, z)$ pure strategies for i's opponents that lead to outcome z:

$$f_i(\theta)[y^i_\tau, r^i(y^i_\tau), z] = \theta^i(y^i_\tau)\sum_{s^i \in \hat{s}^{-i}(s^i,z)} \prod_{j \neq i} \bar{\theta}^j(s^j).$$

where $f_i(\theta)[y^i_\tau, r^i(y^i_\tau), z] = 0$ if $s^i \neq r^i(y^i_\tau)$.

new entrants to the population have no experience: $f^i(\theta)[y^i_0] = 1/T$.

The Dynamic

$\theta_{t+1} = f(\theta_t)$

steady state $\theta = f(\theta)$

Steady State in Strategy Space

given $\bar{\theta}$, we may assume that you meet people drawn from that distribution over your life, and starting with $\tilde{f}_i(\bar{\theta})[y_0^i] = 1/T$ we can work out recursively

$$\tilde{f}_i(\bar{\theta})[y_\tau^i, r^i(y_\tau^i), z] = \tilde{f}_i(y_\tau^i) \sum_{s^i \in \hat{s}^{-i}(s^i, z)} \prod_{j \neq i} \bar{\theta}^j(s^j)$$

then aggregate

$$\bar{f}^i(s^i) \equiv \sum_{y_\tau^i | r^i(y_\tau^i) = s^i} \tilde{f}_i(y_\tau^i)$$

a steady state is also characterized by $\bar{\theta} = \bar{f}(\bar{\theta})$

note that the dimensionality here is independent of T

Patient Stability

a sequence of steady states $\lim_{T\to\infty} \overline{\theta}^T \to \overline{\theta}$ we say that $\overline{\theta}$ is a g^0, δ-*stable state*

If $\overline{\theta}(\delta)$ are g^0, δ-*stable states* and $\lim_{\delta\to 1} \overline{\theta}(\delta) \to \overline{\theta}$, we say that $\overline{\theta}$ is a *patiently stable state.*

Theorem 5.1: *(Fudenberg and Levine [1993])* g^0, δ-steady states are self-confirming; patiently stable states are Nash.

The Converse?

Does patient stability do more than Nash?

Looks as though it requires off-path experimentation – could this give subgame perfection?

Simple Games

a *simple game*

- ♦ perfect information (each information set is a singleton node)
- ♦ each player has at most one information set on each path through the tree. (may have more than one information set, but once he has moved, he never gets to move again)

generic condition: no own ties

weaker than no ties

Subgame Confirmed Nash Equilibrium

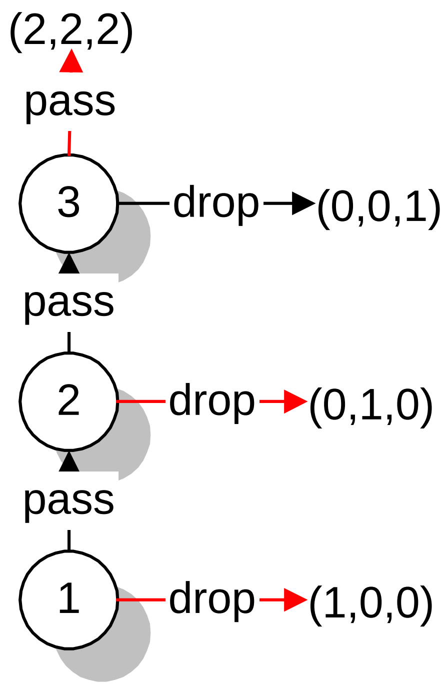
In a simple game, node x is one step off the path of π if it is an immediate successor of a node that is reached with positive probability under π.

Definition: *Profile* π *is a* subgame-confirmed Nash equilibrium *if it is a Nash equilibrium and if, in each subgame beginning one step off the path, the restriction of* π *to the subgame is self-confirming in that subgame.*

In a simple game with no more than two consecutive moves, self-confirming equilibrium for any player moving second implies optimal play by that player, so subgame-confirmed Nash equilibrium implies subgame perfection.

can fail when there are three consecutive moves.

Example: The Three Player Centipede Game



unique subgame-perfect equilibrium: all players to **pass**

(**drop**, **drop**, **pass)** is subgame-confirmed

The Problem of Hammurabi

"If any one bring an accusation against a man, and the accused go to the river and leap into the river, if he sink in the river his accuser shall take possession of his house. But if the river prove that the accused is not guilty, and he escape unhurt, then he who had brought the accusation shall be put to death, while he who leaped into the river shall take possession of the house that had belonged to his accuser." [2nd law of Hammurabi] puzzling to modern sensibilities for two reasons

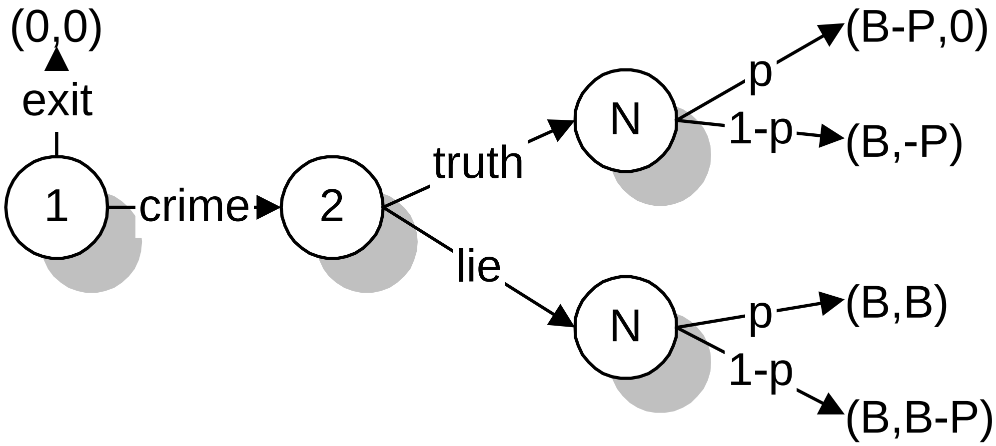
- based on a superstition that we do not believe to be true – we do not believe that the guilty are any more likely to drown than the innocent
- if people can be easily persuaded to hold a superstitious belief, why such an elaborate mechanism? Why not simply assert that those who are guilty will be struck dead by lightning?

from the perspective of the theory of learning in games we ask: which superstitions survive?

- Hammurabi had it exactly right: (our simplified interpretation of) his law uses the greatest amount of superstition consistent with patient rational learning

The Hammurabi Games

Example: The Hammurabi Game

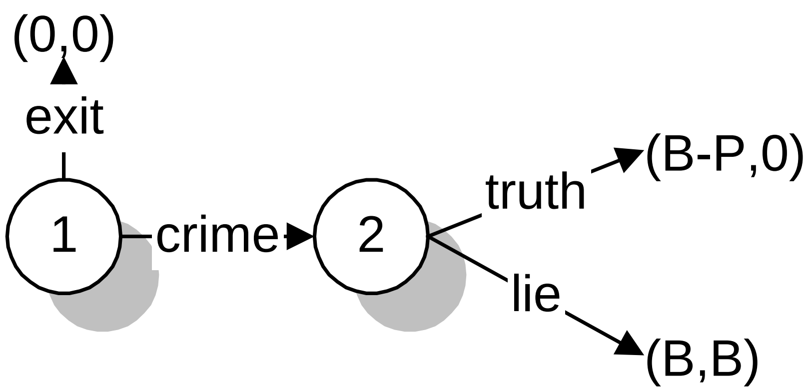


loosely inspired by the law of Hammurabi; player 1 is a suspect; player 2 an accuser; everyone knows the crime has taken place; abstracts from the death penalty

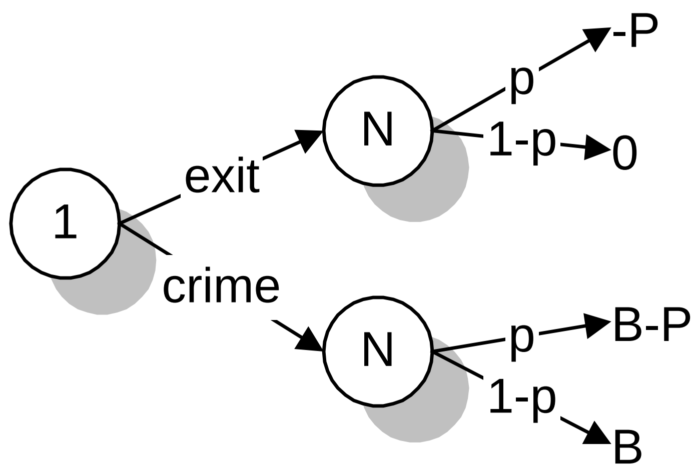
B is the benefit to the accuser of a **lie**, to the suspect of **crime**

P is the loss being punished; probability of punishment sufficient to deter **crime**, $B < pP$

Example: The Hammurabi Game Without a River



Example: The Lightning Game



configurations in which there is no crime

Hammurabi game (Nash, but wrong beliefs about off-off path play)

♦ accuser tells the **truth** because he believes that if he **lie**s he will be punished with probability 1

Hammurabi game without a river (Nash, but not off-path rational)

♦ accuser tells the **truth,** and is indifferent (ex ante, not ex post)

lightning game (self-confirming, but not Nash)

♦ everyone believes that if they commit a **crime** they will be punished with probability 1, and that if they **exit** they will be punished with probability p

All the Hammurabi games are simple games

Patient Stability in Simple Games

two profiles $\overline{\theta}, \overline{\theta}'$ are *path equivalent* if they induce the same distribution over terminal nodes.

a profile is *nearly pure* if Nature does not randomize on the equilibrium path, and no player except Nature randomizes off the equilibrium path

our proposed Hammurabi game profile is nearly pure – only Nature randomizes, and only off the equilibrium path

Theorem: *In simple games with no own ties, a subgame-confirmed Nash equilibrium that is nearly pure is path equivalent to a patiently stable state.*

♦ randomization by players off the equilibrium path – can accomplish this through purification and types

♦ randomization by Nature on the equilibrium path – in an infinite horizon discounted one-armed bandit problem does the probability of getting stuck on the wrong arm go down at the rate $(1-\delta)^2$ or faster?

Necessity of subgame-confirmed: affirmative with "independent beliefs" (not in paper)

♦ without independent beliefs it may be desirable at an off path node to experiment to generate information about an on path node

Definition*: A profile π is ultimately admissable if no weakly dominated strategy (action) is played in an ultimate subgame.*

Remark: every subgame confirmed Nash equilibrium is ultimately admissable. In a simple game with no more than two consecutive moves, Nash equilibrium plus ultimate admissability is equivalent to subgame perfection, hence to subgame confirmed Nash equilibrium.

Theorem: *Patiently stable states are ultimately admissable Nash equilibria.*

This answers the Hammurabi puzzle: the Hammurabi equilibrium with the river is patiently stable; without the river it is not ultimately admissable; lightning equilibrium even Nash

Games with Length at Most Three

a game has "length at most three" if no path through the tree hits more than three information sets

Theorem *In simple games with no own ties, no Nature's move and length at most three, a subgame-confirmed Nash equilibrium is path equivalent to a patiently stable state.*

because in these games all equilibria are nearly pure

Lemma: *In simple games with no own ties, no Nature's move and length at most three, a subgame-confirmed Nash equilibrium is path equivalent to a subgame-confirmed Nash equilibrium in which players play pure strategies.*

in turn follows from

Lemma: *In simple games with no own ties, no Nature's move and length at most two, every self confirming equilibrium is path equivalent to a public randomization over Nash equilibria.*