Chaos, Complexity, and Inference (36-462)

Lecture 25: Adaptive Behavior

Cosma Shalizi

21 April 2009

The Dollar Auction

Adaptive Behavior

Games

Evolutionary Games

Reinforcement

Cascades

Networks from Games

The best introductory textbook on game theory is Gintis (2000). Less technical but good orientations: Poundstone (1992); Sigmund (1996); Slee (2006)

Games

Agents or players; "Nature" may be a player

Actions or moves

Pay-off reward or punishment for each player's action, *given* all the others' moves

Single-valued utilities are actually a very dubious assumption, on basic neurological grounds (McCulloch, 1945)

Game tree shows history of moves by all players to date

Strategy says which move to make at each node in game tree (possibly stochastic)

Best reply move/strategy which has highest pay-off *given* other player's moves/strategies

Equilibrium everyone plays best reply against everyone else

Dominated strategy Another strategy *always* does at least as well, and sometimes better

Minimax minimize the maximum harm suffered

"Rational" maximizing subjectively-expected payoffs, with personal, subjective probabilities updated by Bayes's rule; sometimes with extra assumption that subjective expectations are always objectively unbiased

"rationality" $\Rightarrow$ "elimination of dominated strategies"

Backwards induction Recursive elimination of dominated strategies

"Rationality" in action (1): the ultimatum game

"Rationality" in action (2): the prisoners' dilemma

Bounded rationality: not fully "rational", but uses an actual, implementable procedure to make decisions (Simon, 1955, 1956)

Institutions simplify decisions so people can make them

Evolutionary Games

The classic work: Maynard Smith (1982)
Pay-offs are to strategies, which are the **replicators**; pay-off is now **fitness**, $f(s)$
Dynamics concern the population share or frequency $p(s)$ of the *replicators*
higher fitness $\Rightarrow$ bigger population share
implementations: genetics, imitation

Replicator Dynamics

Replicator equation:

$$p_{t+1}(s) = p_t(s)\left[f_t(s) - \sum_{s'} f_t(s')p_t(s')\right]$$

Note that

$$\sum_s \Delta p_t(s) = 0$$

so it stays normalized, $\sum_s p_t(s) = 1$
Defines a dynamical system which we can analyze like any other (Hofbauer and Sigmund, 1998)

ESS

Evolutionarily stable strategy: one which can't be invaded

For any $s' \neq s$, $\Delta p(s') < 0$ when $p(s) = 1 - \epsilon$, $p(s') = \epsilon$, ϵ sufficiently small

Not all equilbria are evolutionarily stable!

ESS = stable fixed point

Cooperation in Prisoner's Dilemma

Play with automata
Always cooperate vs. always defect: defect wins
Tit-for-tat
Tit-for-two-tats, etc.
Lindgren (1996) summarizes this line of thought
Spatial structure and spatial pattern formation: nice discussion in Sigmund (1996)

Reinforcement

Adaptation within individual, not across population
"weight" of action s

$$
\begin{aligned}
w_{t+1}(s) &= \alpha w_t(s) + (1-\alpha) f(s) \text{ if played } s \\
w_{t+1}(s) &= \alpha w_t(s) \text{ otherwise} \\
p_t(s) &= \frac{w_t(s)}{\sum_{s'} w_t(s')} \text{ or} \\
p_t(s) &= \frac{\exp w_t(s)}{\sum_{s'} \exp w_t(s')}
\end{aligned}
$$

Can do likelihood inference since this gives probabilities for observable actions

Notes on reinforcement:

1. "strategies" here do not have to be single moves but could be complicated
2. can give similar dynamics to replicator equation (Börgers and Sarin, 1997; Borkar, 2002; Sato and Crutchfield, 2003)
3. Many variants on shape of the reinforcement, precise learning dynamics, etc. — Sutton and Barto (1998) analyzes many versions used in AI and robotics
4. Experimentally, reinforcement learning can give *excellent* matches to human data (Salmon, 2001; Erev and Roth, 2001)
5. RL is close to "multiplicative weight training" in machine learning, which leads to low **regret** = difference between actual payoff and payoff of best single strategy

Normal human beings seem regret-driven (Marchiori and Warglien, 2008), but not those with orbifrontal lesions (Camille *et al.*, 2004)

Convergence via Reinforcement

Polya's urn: start with one ball of each of k colors
X_t = color of ball drawn from urn, uniformly, at time t
put that ball back, and add another of that color

$$p_{t+1}(s) = \frac{p_t(s)(k+t) + \mathbf{1}_s(X_t)}{k+t+1}$$

Analysis of the urn model:

$$\begin{aligned} \mathbf{E}\left[p_{t+1}(s)\right] &= \frac{k+t}{k+t+1} p_t(s) + \frac{1}{k+t+1} \mathbf{E}\left[\mathbf{1}_s(X_t)\right] \\ &= \frac{k+t}{k+t+1} p_t(s) + \frac{p_t(s)}{k+t+1} \\ &= \frac{k+t+1}{k+t+1} p_t(s) = p_t(s) \end{aligned}$$

so $p_t(s)$ is a **martingale**
Bounded martingales converge almost surely $\Rightarrow$ p_t converges a.s.

General flavor of analysis holds much more generally: under reasonable conditions, if

$$\mathbf{E}\left[p_{t+1}\right] = f_n(p_t)$$

and

$$f_n \rightarrow f$$

then long-run behavior of p_t tracks that of the *deterministic* dynamical system

$$x_{t+1} = f(x_t)$$

(Arthur, 1994; Pemantle, 2007)

Cascades

Self-reinforcing actions in games

- Information cascades
- Coordination

Experiment: Salganik *et al.* (2006)

Networks and Games

Play with your neighbors: similar effects to spatial structure
actually, discrete space is a special case
Skyrms and Pemantle (2000); Pemantle and Skyrms (2004): two decisions, who to play with and what strategy to follow
Reinforce ties tha lead to good pay-offs
Leads to *endogeneous* network formation

Arthur, W. Brian (1994). *Increasing Returns and Path Dependence in the Economy*. Economics, Cognition and Society. Ann Arbor: University of Michigan Press.

Börgers, Tilman and Rajiv Sarin (1997). "Learning Through Reinforcement and Replicator Dynamics." *Journal of Economic Theory*, **77**: 1–14. doi:10.1006/jeth.1997.2319.

Borkar, Vivek S. (2002). "Reinforcement Learning in Markovian Evolutionary Games." *Advances in Complex Systems*, **5**: 55–72. doi:10.1142/S0219525902000535.

Camille, Nathalie, Giorgio Coricelli, Jerome Sallet, Pascale Pradat-Diehl, Jean-René Duhamel and Angela Sirigu (2004). "The Involvement of the Orbitofrontal Cortex in the Experience of Regret." *Science*, **305**: 1167–1170. doi:http://dx.doi.org/10.1126/science.1094550.

Erev, Ido and Alvin E. Roth (2001). "Simple Reinforcement Learning Models and Reciprocation in the Prisoner's Dilemma Game." In *Bounded Rationality: The Adaptive Toolbox* (Gerd Gigerenzer and Reinhard Selten, eds.), vol. 84 of *The Dahlem Workshops*, pp. 215–232. Cambridge, Massachuetts: MIT Press.

Gintis, Herbert (2000). *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton: Princeton University Press.

Hofbauer, Josef and Karl Sigmund (1998). *Evolutionary Games and Population Dynamics*. Cambridge, England: Cambridge University Press.

Lindgren, Kristian (1996). *Evolutionary Dynamics in Game-Theoretic Models*. Tech. Rep. 96-06-043, Santa Fe

Institute. URL http://www.santafe.edu/research/publications/wpabstract/199606043.

Marchiori, Davide and Massimo Warglien (2008). "Predicting Human Interactive Learning by Regret-Driven Neural Networks." *Science*, **319**: 1111–1113. doi:10.1126/science.1151185.

Maynard Smith, John (1982). *Evolution and the Theory of Games*. Cambridge, England: Cambridge University Press.

McCulloch, Warren S. (1945). "A Heterarchy of Values Determined by the Topology of Nervous Nets." *Bulletin of Mathematical Biophysics*, **7**: 89–93. Reprinted in (McCulloch, 1965, pp. 40–45).

— (1965). *Embodiments of Mind*. Cambridge, Massachusetts: MIT Press.

Pemantle, Robin (2007). “A Survey of Random Processes with Reinforcement.” *Probability Surveys*, **4**: 1–79. URL `http://arxiv.org/abs/math.PR/0610076`.

Pemantle, Robin and Brian Skyrms (2004). “Network formation by reinforcement learning: the long and medium run.” *Mathematical Social Sciences*, **48**: 315–327. URL `http://arxiv.org/abs/math.PR/0404106`.

Poundstone, William (1992). *Prisoner’s Dilemma*. New York: Doubleday.

Salganik, Matthew J., Peter S. Dodds and Duncan J. Watts (2006). “Experimental study of inequality and unpredictability in an artificial cultural market.” *Science*, **311**: 854–856. URL `http://www.princeton.edu/~mjs3/musiclab.shtml`.

Salmon, Timothy C. (2001). "An Evaluation of Econometric Models of Adaptive Learning." *Econometrica*, **69**: 1597–1628.

Sato, Yuzuru and James P. Crutchfield (2003). "Coupled replicator equations for the dynamics of learning in multiagent systems." *Physical Review E*, **67**: 015206. URL `http://arxiv.org/abs/nlin.AO/0204057`.

Sigmund, Karl (1996). *Games of Life: Explorations in Ecology, Evolution and Behavior*. Penguin.

Simon, Herbert A. (1955). "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics*, **69**: 99–118. URL `http://cowles.econ.yale.edu/P/cp/p00b/p0098.pdf`. Reprinted in Simon (1982).

— (1956). "Rational Choice and the Structure of the Environment." *Psychological Review*, **63**: 129–138. Reprinted in Simon (1982).

— (1982). *Models of Bounded Rationality*. Cambridge, Massachuetts: MIT Press.

Skyrms, Brian and Robin Pemantle (2000). "A Dynamic Model of Social Network Formation." *Proceedings of the National Academy of Sciences (USA)*, **97**: 9340–9346. URL `http://arxiv.org/abs/math.PR/0404101`.

Slee, Tom (2006). *No One Makes You Shop at Wal-Mart: The Surprising Deceptions of Individual Choice*. Toronto: Between the Lines.

Sutton, Richard S. and Andrew G. Barto (1998). *Reinforcement Learning: An Introduction*. Cambridge, Massachusetts: MIT

Press. URL `http://www.cs.ualberta.ca/~sutton/book/the-book.html`.