Going Back to El Farol (?)

Michael Grinfeld

Summary

In this talk I will review W. B. Arthur's paradigmatic example of decision-making under uncertainty (Inductive Reasoning and Bounded Rationality, AER **84**(2) (1994), 406–411). In particular, I will examine Arthur's blanket assumptions. The aim is to have a discussion that will encompass affect and memory and will lead to more realistic modelling, in particular, mean-field modelling.

Plan

- Some quotations
- The El-Farol bar set-up
- W. Brian Arthur's solution
- Some other possibilities

Some quotations

. . . beyond a certain level of complexity human logical capacity ceases to cope — human rationality is bounded.

. . . in interactive situations of complication, agents cannot rely on other agents they are dealing with to behave under perfect rationality, and so they are forced to guess their behaviour.

Modern psychologists are in reasonable agreement that in situations that are complicated or ill-defined, humans use characteristic and predictable methods of reasoning. These methods are not deductive but *inductive*.

. . . we look for patterns; and we simplify the problem by using these to construct temporary internal models or hypotheses or *schemata* to work with.

The El Farol bar setup

N people ($N = 100$) decide independently each week whether to go to a bar that offers entertainment on a certain night. Space is limited and the evening is enjoyable if ... fewer than 60 percent of the possible 100 are present. There is no sure way to tell the numbers coming in advance; therefore a person goes ... if he expects fewer than 60 to show up or stays home if he expects more than 60 to go.

Also note:

... any commonality of expectations gets broken up: if all believe few will go, all will go. But that would invalidate this belief.

W. B. Arthur's solution

- Choices are unaffected by previous visits; there is no collusion or prior communication among the agents; the only information is the numbers who came in past weeks.
- Each (simulated) agent is given a set of k focal predictors out of the total set of M (numbers not specified in the AER paper).

  [An example of a predictor is: predict the number next week to be the average of the last 4 weeks.]
- Every week, each agent chooses the predictor that did best last week.

Computer simulations seem to converge to 60% bar occupancy on the average; there are no theorems.

To summarise: in W. B. Arthur's there is no affect; the only memory is of attendance number, and the decision-making process of an agent is the result of competition among randomly assigned predictors chosen from some common pool.

Many questions arise:

- Why model, and why model this situation?
- Do we on some theoretical grounds expect convergence to 60% occupancy on average?
- What do the agents "really" remember?
- What, of what they remember, do they take into account in decision-making?
- How does one take into account both the attendance information and the affective data?

  [By affective data we mean data about categories of psychological experiences, such as Schadenfreude and regret, which agents use to organise their personal experience.]

Some other possibilities

1. In Cross *et al.* (2005) we suggested a purely affective model, in which each agent was described by a vector of psychological tensions, with the decision each week being determined by the dominant tension (for example, regret). Is that an advance on W. B. Arthur?

2. In theory, it is not difficult to formulate *mean field models*, which are in fact amenable to mathematical analysis. Let m_t be the proportion of agents that come to the bar in week t. Then, assuming "minimal" memory,

$$m_{t+1} = m_t P_1(m_t) + (1 - m_t) P_2(m_t). \tag{1}$$

For example, $P_1(s)$ can be a monotone decreasing function of s, with a big drop at $s \approx 0.6$, and $P_2(s)$ can be even taken to be constant. With such simple assumptions, if the constant is well chosen and the function $P_1(s)$ is steep enough, (1) can have chaotic dynamics.

Some more questions

1. What statistics of attendance would count as “rational”?

2. Is it possible to construct a mean-field theory that takes affective data into account?

3. Does taking affective data into account *necessarily* make the agent less rational (by some criterion)?

4. How to model communication/herding effects?

5. What is the empirical data in such situations?

[which of course are relevant in Internet usage etc., “coordination failures”]

6. What happens if there 2, . . . , N such bars?