R. Axelrod – An Evolutionary Approach to Norms

Kian Mintz-Woo
University of Amsterdam

Institute for Logic, Language and Computation

January 22, 2010

INTRODUCTION

- I begin by examining what the author says about the definition of norms. Specifically, Axelrod's approach contrasts with Young's in that norms are not necessarily best responses.
- Furthermore, in Axelrod's system, norms are identified behaviorally as well as incrementally.
- I detail the Norms Game and then the Metanorms Game.
- After doing so, I discuss several of the mechanisms that can be used to reinforce normative behavior, focussing on metanormativity.

OVERVIEW

GOALS FOR MODELING OF NORMS

Before examining the games that Axelrod provides, it is worth looking at what his analysis is intended to show:

1. How norms arise–
2. How norms are maintained–
3. How one norm displaces another–

In other words, “what is needed is a theory that accounts not only for the norms existing at any point in time, but also for how norms change over time” (Axelrod 1986: 1096).

Evolutionary Nature of Norms

- The intention of analysis is to find what initial conditions contribute to the growth and success of norms in game-theoretic terms.
- Since norms are approached through iterated games, and through iteration can we determine how norms arise and displace each other, it is natural to approach norms through an evolutionary approach.
- This incremental approach shows how subsequent generations gain from the improvements of previous ones, as well as accounting for some mutation (*i.e.* variation, experimentation).

Behavioral Definition of Norms

Since we are trying to model norms in actual social (*e.g.* political) institutions, a definition which admits of gradations is valuable:

Definition (Norm)

A *norm* exists in a given social setting to the extent that individuals usually act in a certain way and are often punished when seen not to be acting in this way.

The gradations come in two forms in the Norm Game: boldness (*i.e.* how often the action is taken) and vengefulness (*i.e.* how often someone is punished for not following a norm).
NB: The focus is on punishment, not on Nash equilibria.

SETUP

- The population N of size $|N| = n$. Each individual $i \in N$ has a level of Boldness B_i and a level of Vengeance V_i, both on a scale of $0 - 7$.[1]
-

Outcome	Value
Temptation	3
Hurt	−1
Punished	−9
Enforcement	−2

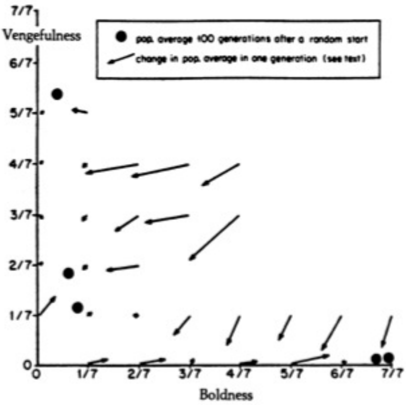
- After acting, an agent has 0.5 chance of being Seen (one assumption is that every player has equal chance of being seen). If a player i has a boldness greater than the chance they'll be seen, *i.e.* $B_i > S$, then that agent defects.

[1] This can be expressed in three bits (!) Computation is truly amazing.

Simulating the Norms Game

1. Initial Strategies: Twenty random strategies (*i.e.* levels of B_i and V_i) are determined from the complete strategy space.
2. One Generation: Each player i is given four chances to defect. For each i, each defection occurs if $B_i > S$. Each defection, each other player j has S chance of noticing, and will P with frequency V_j. The total outcomes are totaled for each player.
3. Evolutionary Stage: For those who do well (*i.e.* one SD above average), they get two offspring. Those who do poorly (*i.e.* one SD below average) get no offspring. Others get one. Then, 1% bit changing for mutation (*i.e.* this occurs for about 1.2% of the succeeding generation bits).
4. A Run: Repeat Steps 2 and 3 for one hundred generations.
5. Ad Nauseum: Repeat Steps 1 to 4 for five complete runs .

A Sample Simulation

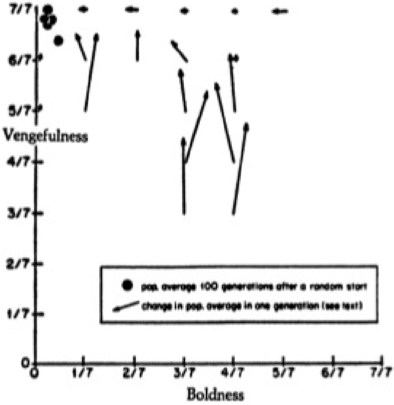


- Initially, average V_i and $B_i \sim \frac{3.5}{7}$.
- Then $\Leftarrow B$, then $\Leftarrow V$ then $B \Rightarrow$ again.
- Dark spots indicate after 100 generations, (1 x) high V, low B, (2 x) low V, low B, (2 x) low V, high B.

SIMULATING THE METANORMS GAME

- Now, we expand the Norms Game to the Metanorms Game, by tweaking the Evolutionary Stage.
- For each of the four possible defections of player i, there is S chance of being seen for any given other player j. If j does S and does not P, then there is S chance of that non-P being S by another player k.
- Each player k who S j will punish j with frequency V_k. The P' and E' costs are the same as P and E (is this reasonable?). *I.e.* , $P' = P = -9$ and $E' = E = -2$.

A SAMPLE SIMULATION



- Initially, average V_i and $B_i \sim \frac{3.5}{7}$.
- We see in few generations, $V \Rightarrow$, then $\Leftarrow B$.
- Dark spots indicate after 100 generations, all outcomes have high V, low B.

METANORMS

- Since failing to punish is as costly as defecting, having low V is costly. Having high V leads quickly to low B. This encourages cooperative behavior.
- The key is linking the vengeance with metavengeance. Does this seem likely? Axelrod says that he suspects there is a correlation in metanormativity, although it is not clear.
- My own conjecture is that social pressures suggest vengeance is linked with metavengeance, but that metavengeance carries significantly weaker P costs than vengeance.
- *E.g.* Consider a group situation where someone does something wrong (*e.g.* in an office someone steals office goods). It seems to me that your vengeance level against that individual would be higher than against someone who saw (and did not report) these infractions (*i.e.* your metavengeance). (Dis)agree?

OTHER MECHANISMS: DOMINANCE

- Dominance occurs when we partition our population into different groups and adjust the payoffs so they are no longer equal.
- Both groups still require metanorms to enforce conventionality, but according to the paper, it is easier for the stronger group.
- Why?
- The P costs for the weaker groups are higher than those of the stronger groups. If the P-cost is low for a strong group, metavengeance carries less weight.

Other Mechanisms: Internalization

- Internalization makes the T lower, since the value of defecting is offset (at least partially) by the psychological cost of defecting.
- This might also be plausibly modeled as greater gain in E.
- When norms are internalized boldness decreases, subsequently vengeance decreases.

OTHER MECHANISMS: LAW

- The connection between law and social norms is not complex: Codification of social norms into legal code (at any level) requires a certain level of societal acceptance.
- Once a law has been ratified, this publicly reinforces the normative behavior.
- Axelrod suggests several ways in which the strength of law supports a burgeoning norm:
  1. It adds state punishment to existing private enforcement mechanisms.
  2. The law has more prima facie effect on people regardless of enforcement.
  3. When norms are codified, they become substantially clearer.

DISCUSSION QUESTIONS

- If a norm becomes codified in law, is it unlikely to go *beyond* law? (Axelrod 1986: 1107)
- Does metanormative punishment/vengeance have the same costs (*I.e.* P'/E') as normative P, E?
- Can vengeance give a positive value with psychological reinforcement? (*E.g.* tattle-telling, nagging, costly enforcer)
- Has Axelrod succeeded in his goals?
  1. How norms arise–
  2. How norms are maintained–
  3. How one norm displaces another–
- Are the values too high? In particular, I think that S, at 0.5, is very significant—If there are 20 agents in a population, does each of them have 50% chance of seeing any particular behavior?