

Plans for Model Extensions

August 18, 2014

In the basic model, there are three interaction types: cooperators, defectors, and discriminators. The discriminators gather, store, and use information about their peers. The collection of information is parameterized by the frequency of observation, f_o . Previously, the storage of information was parameterized by the probability of remembering, p_r . We could also parametrize it with a memory loss parameter dictating the rate at which memories decay. In the previous model, discriminators used information about their peers by cooperating with agents they had seen cooperate and defecting against agents they had seen defect, first order indirect reciprocity. We could also consider second order norms. These rules specify how to interact with an agent based on both his observed action and the identity of his past interaction partner. As such, they can be described by a 2×2 matrix with the rows representing the behavior of the agent and the columns representing the type of his interaction partner and where the entries show the effect of the behaviors on the reputation of the agent:

	C	D
C	+1	j
D	-1	+1

There are 2^4 possible social norms, but there are two popular variants, stern and mild judging where $j = -1$ and $j = 1$ respectively. We could allow a discriminator to choose between these alternatives with a probability p_s . We could also consider the so-called “leading eight” social norms and consider them as discrete discriminator types, perhaps considering group selection between groups using different social norms. To summarize, there are three dimensions of information behavior:

1. collection, parameterized by f_o
2. storage, parametrized by p_r or memory loss
3. usage, parametrized by p_s

In previous work, I derived the following payoff for each of the three interaction types, which depend on the information behaviors of the discriminators:

$$\begin{aligned}
 P_{\text{def}} &= bRx_{\text{coop}} + p_sb(R - \bar{K})x_{\text{disc}} \\
 P_{\text{coop}} &= P_{\text{def}} - cR + b\bar{K}x_{\text{disc}} \\
 P_{\text{disc}} &= P_{\text{def}} - c(R - \bar{K})p_s - c\bar{K}x_{\text{coop}} + (b - c)\bar{G}x_{\text{disc}} - s
 \end{aligned}$$

If we subtract P_{def} from each payoff function we get

$$\begin{aligned}
 P_{\text{coop}} &= -cR + b\bar{K}x_{\text{disc}} \\
 P_{\text{def}} &= 0 \\
 P_{\text{disc}} &= -c(R - \bar{K})p_s - c\bar{K}x_{\text{coop}} + (b - c)\bar{G}x_{\text{disc}} - s \\
 \bar{P} &= x_{\text{coop}}P_{\text{coop}} + x_{\text{def}}P_{\text{def}} + x_{\text{disc}}P_{\text{disc}}
 \end{aligned}$$

The replicator equations describing how the frequencies of the three interaction types change over time are given by

$$\dot{x}_i = x_i(P_i - \bar{P}). \quad (1)$$

I will use fast-slow dynamics to model the evolution of information behaviors. Let $s_r(m)$ denote the growth rate of a mutant discriminator with trait m in a resident population of discriminators with trait r . The following equation describes how the information trait y will change over time:

$$\lambda \frac{dy}{dt}(x_1, x_2, x_3, y) = \left. \frac{\partial s_y}{\partial m} \right|_{m=y} \quad (2)$$

This fast-slow formalism allows us to impose a timescale separation between the replicator dynamics and the adaptive dynamics and the two extreme cases, $\lambda = 0$ and $\lambda = \infty$, can be studied analytically. If $\lambda = \infty$, then \vec{x} changes infinitely quickly with respect to the trait y , for instance if the replicator dynamics happen on ecological timescales and y changes on an evolutionary timescale. If $\lambda = 0$, then x changes infinitely quickly with respect to the frequencies \vec{x} , for instance if discriminators can learn how to use information given the environment set by the type frequencies. For $\lambda = 0$, I will identify the critical manifold $x^* = C(x_1, x_2, x_3)$ such that y^* is a stable equilibrium if x_1, x_2, x_3 are fixed. Once y reaches this critical value, the replicator dynamics give the gradient describing how the type frequencies should change. The converse of this procedure will show us how y evolves if it changes more slowly than the type frequencies. I can also study an intermediate situation where the type frequencies and the trait change on similar frequencies, which would require numerical analysis.

This model will allow us to describe the evolution of not just one but multiple information traits together. Having multiple traits changing will require us to consider

- the cost of increasing information gathering, storage, and more complicated processing
- tradeoffs between the traits, e.g. if making more observations means you get to interact less frequently or off making more observations also tends to increase memory

Big questions:

1. At equilibrium, what information behaviors strategy will discriminators use?
2. Does allowing the information traits to evolve stabilize cooperation, destabilize cooperation, or neither?
3. How does this depend on the timescale separation, i.e. whether information traits change more or less quickly than the type frequencies?

Steps:

1. evolution of f_o under extreme cases $\lambda = 0$ and $\lambda = \infty$
2. evolution of f_o at intermediate timescale separation
3. evolution of f_o and p_s under extreme cases
4. group selection between groups with discriminators using different social norms

Points to keep in mind:

- read up on trait tradeoff analysis [1], critical function analysis, and evolutionary branching
- it seems likely that we will observe oscillations in the amount of information discriminators have and the frequency of cooperators as a high frequency of cooperators will make information superfluous and when information decreases defectors will be able to invade more easily
- it's been observed previously that the Kandori social norms requires discriminators to have a lot of information so we should keep in mind whether our results corroborate or question this claim

References

- [1] Claire deMazancourt and Ulf Dieckmann. Trade-off geometries and frequency-dependent selection. *The American Naturalist*, 164(6), 2004.