Inter-Group Empathy: Simulation Outline

Gavin Rees

June 2019

1 Overview

The simulation is based on a finite population of type A and type B agents that play a donation game with each other, all under a social norm based reputation system. Each generation entails every agent playing the donation game with every other agent. All the agents observe one of each of the other agents' interactions per generation and judge them based on the social norm, storing these new reputations privately. With some probability ($E_{i,A}, E_{i,B}$ for intra-group judgement, $E_{o,A}, E_{o,B}$ for inter-group judgement) this judgement is empathetic, meaning the judge uses the agent's views of each other's reputations to judge their actions, as opposed to using judge's own subjective reputation list to judge them.

To reduce the complexity of the model, the agents only have strategies of the types AllC, Disc, and AllD: cooperate with all individuals, cooperate with "good" individuals and defect for "bad" individuals, and defect for all individuals, respectively.

Every generation the population is updated according to the model of social imitation learning. Two agents are selected at random from the population, and the first (arbitrarily chosen) takes on the second's type and strategy with probability $\frac{1}{1+e^{-\omega(\pi_2-\pi_1)}}$ where π_1 is the first agent's payoff from that generation, and π_2 the second's. There's also a small probability, μ, that a given agent's strategy randomly changes after each round.

Note: right now empathy parameters are globally fixed, not varying throughout the population or changing over the course of the simulation.

2   Parameters

- Norm: represented as a matrix with the row indices representing the agent's action, and the column indices representing the recipients reputation. Note that, for actions, 1 represents cooperate and 0 represents defect, and for reputations 1 represents "good" and 0 represents "bad".
  Current norm is stern judging

- Error rates: e_1 the probability that intended cooperation switches to defection, and e_2 the probability that incorrect reputation is assigned.
  Currently $e_1 = 0, e_2 = 0$

- Empathy: E the probability that a judgement is empathetic.
  Currently $E = 0$

- Population size, N
  Currently $N = 100$

- Initial proportion of type A, $f_{A,0}$
  Currently $f_{A,0} = 0.5$

- Number of generations
  Currently 1500, same as original paper

- Social learning rate ω
  Currently $\omega = 1$, same as original paper

- Random exploration rate, μ
  Currently $\mu = 0.0025$, one of several tested in original paper

3   Things I'm not sure about:

- Does it make sense to use the update rule derived for social imitation learning to examine updating in intergroup dynamics? Would death-birth updating make more sense? What choice would make it easiest to later model this with a replicator-dynamics sort of ODE?

- Does imitating empathy happen in an absolute sense (imitate their empathy towards type A and their empathy towards type B) or in a relative sense (imitate their in-group and out-group empathy, applied

to the imitators particular type). This is another reason I'm debating a different update rule.

- What's the least complicated way for random exploration rate to effect type vs. strategy?

- What are some interesting/informative ways to initialize all of the parameters that vary throughout the population?

4 Some initial questions I'm interested in exploring

- What happens with one group that's internally completely empathetic, and another group that has 0 empathy?

- What about entirely out-group empathy, no in group empathy?

- What's the time scale of an fixation of type A or B?

- What sort of subjective reputations do members of one group form for members of the other group? How does this change as inter-group and intra-group empathy varies?