

CSCI 8551 Intelligent Agents Course Project

Final report

Refined Model in Python: the Thriving of Punishers

Bolei Di, Xi Guo

December 17, 2015

1 Introduction

Punishment in a society allowing voluntary cooperation has been evolving for more than two thousand years [1]. Public punishment by paying a personal cost has proved to be an important mechanism that facilitates the cooperation of humans [2]. In the study of Hauert *et al.*, a model has been exploited to show the evolutionary merits of punishment. Emergence of punishers can be achieved when joining the cooperation is voluntary. The study of Hauert *et al.* was based on several assumptions that may not represent the characteristics of human organizations. Specifically, the original model were created in an ideal environment in which each agent has the full access to the performance of other agents. The agents were also not obligated in contributing to the punishment. However, we thought of two examples that contradicts these assumptions in the model. 1) The restricted communication infrastructure in ancient time (such as beacon) provided incomplete information, which people made their decisions based upon. 2) One of the functions of the obligated tax system is to establish the institutes for enforced punishment, including police and jails. To address these limitations in the original model, in this study, we attempted to rebuild the model and enhance the model in order to explore the theoretical aspects of the situations described above.

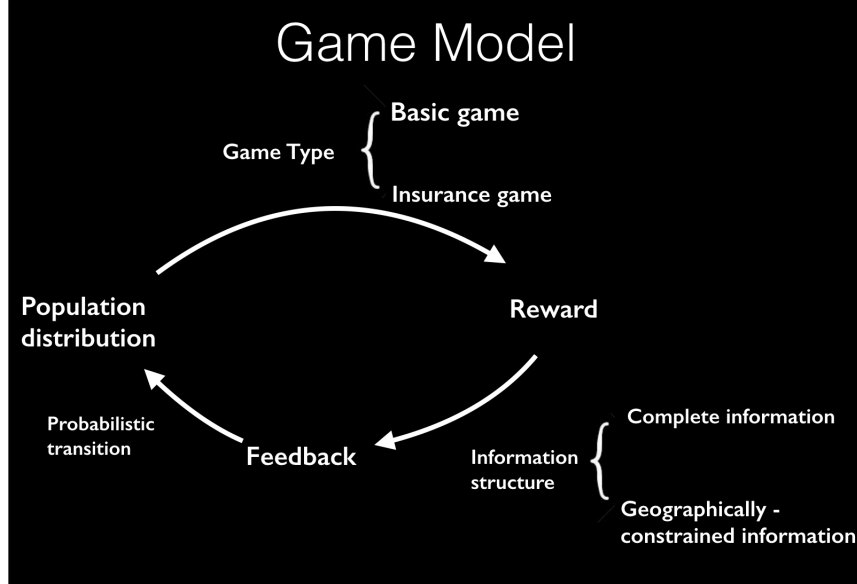
We planned to add the following variations in the new model. 1) Geographical constraints of the agents. 2) Enforced insurance (or tax) policy for each agents in the society. 3) Combining the geographical constraints and the insurance policy. We proposed to achieve three goals: 1) To replicate the original result; 2) To build enhanced model by incorporating information constraints and the enforced insurance policy; 3) To examine the impact of variations and parameters on the current model.

2 Game Model

The game model consists of three parts which forms a loop that enables running the model for infinite steps as in Figure 1. Game type determines the reward to each type of agents given a

population of agents. Information structure determines the feedback each agent sees based on rewards. Transition probability takes in the feedback of each agent and output its probabilities of transforming into different types of agents. Given each different method for each step, this model can simulate any combination of the three steps.

Figure 1: Overview of game model



2.1 Game type

Here we consider two game types for the experiments - basic game and enforced insurance game. We use $R_n, R_c, R_p, R_d, n, c, p, d$ to denote the rewards and populations for non-participant, cooperator, punisher and defector. l denotes the contribution of cooperators and punishers' contribution to the cooperation. γ denotes the cost of punishment for each punisher and β denotes the punishment a defector receive from each punishment. t is the tax or insurance paid by each member in the cooperation in insurance game. Rewards of a basic and insurance game is defined as:

$$s = c + p + d \quad (2.1)$$

$$R_n = \sigma \quad (2.2)$$

$$R_c = r(c + p)l/s - l - \{t\} \quad (2.3)$$

$$R_p = r(c + p)l/s - l - d\gamma - \{t\} \quad (2.4)$$

$$R_d = r(c + p)l/s - p\beta + \{t(c + p)/d\} \quad (2.5)$$

where the terms in $\{\}$ are included only in insurance game. All the results shown later in this report share the same set of parameter values: $\sigma = 1, r = 5, l = 1, \beta = 2, \gamma = 0.1, t = 0.1$. These parameters are carefully tuned, such that, 1) having one punisher in the cooperation, cooperators' rewards would be higher than defectors, 2) non-participants receive much lower (one fourth) rewards than anyone in a healthy (few defectors) cooperation, 3) a punishment is very cost-effective and 4) a tax/insurance paid by members is only a small fraction of rewards in healthy cooperation.

2.2 Information structure

For the basic game, the above information is seen by all agents, and for the geographically constrained game, an agent only sees its neighbors' rewards and only knows about types of agents of itself and its neighbors. For our geographically constrained simulations, the game is played in a 50 by 50 rectangle region and the visibility radius is 10.

2.3 Probabilistic transition

The transition function determines the probabilities of transforming into other agent types given the feedback an agent sees. We have several transition function but only one worked out well. The transition function distinguishes between two states of agents, ignorance and non-ignorance. An ignorance agent sees only type and rewards of its own kind, has 99.7% chance staying as is and 0.3% chance transforming into other agents. For a non-ignorance agent, it still has 80% chance of staying as is, 1% for each type it does not see, and the rest close to 20% is splitted proportionally to the sigmoid function of the differences between its own reward and other agent types' reward it can see.

3 Experimental Results

In the following sections, we will illustrate the results from the basic game, enforced insurance game and geographically constrained game. We will discuss the effect of these two features.

3.1 Basic game

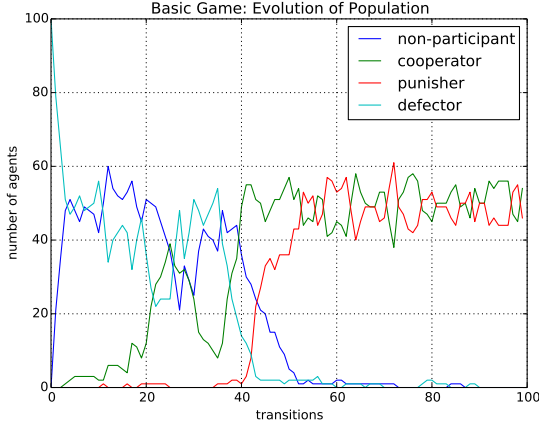
As shown in Figure 1(a) and Figure 1(b), the basic game is initialized with 100 defectors. (Other parameters need to be listed here) We observed the similar trend of the successful evolution of punishers after 100 transitions. Both cooperators and punishers have survived in the end and maintained in a stable state. Unlike Figure 1A in the original paper, we failed to observe the dominant trend of a overly large number of punishers even after 1000 transitions (Figure 1(c)). The possible explanation for the distinction of the number of punishers is that the transition function that we used allows a smaller number of emerging defectors, thus the system does not require a large number of punishers to reach to the steady state. Furthermore, we have observed the similar oscillation transient of non-participants, cooperators and defectors. This transient period is from the starting point to the point when the system reaches to a stable state.

3.2 Effect of geographic constraints

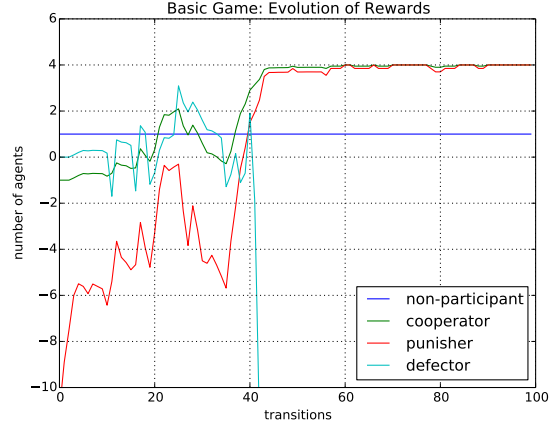
Geographic constraint worked as a inhibitor in the process of evolution. It increased the oscillation of the number of agents and sometimes it could slow down the propagation of one or more types of rewarding agents. Comparing Figure 1(a) and Figure 2(a), we discovered that same initialization condition has led to almost the same lengths of the transient period (about 60 transitions). The population distribution exhibited higher fluctuation in the geographically constrained game. The

Figure 2: Basic game

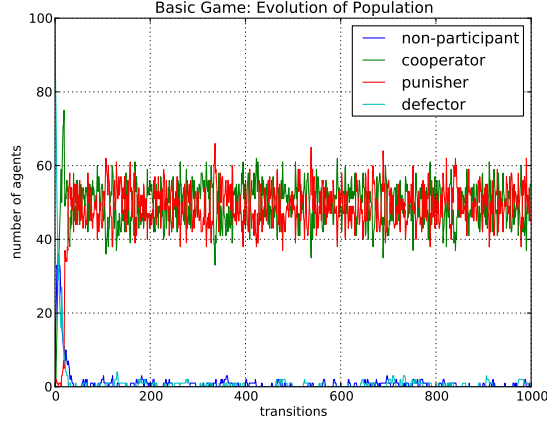
(a) The evolution of agents: basic game



(b) The Evolution of rewards: basic game



(c) Results of basic game (1000 transitions)



constrain could further slow down the propagation of one or more types of rewarding agents, as its good performance was not seen by its neighbor agents. Figure 3(a) and Figure 3(b) are two snapshots of the distribution of agents in the game corresponding with Figure 2. The gray circle indicates an agent has been transformed to punisher at 17th iteration. Although the reward of punishers at this point was as high as that of defectors', being restricted by the geographic constraint, the punisher was not able to influence its neighboring agents. It changed to a defector in the subsequent transition.

3.3 Effect of insurance

As seen in Figure 4, the enforced insurance policy worked as a catalyst for the evolution by keeping the system stable. First of all, it made the transient period shorter by changing it from 60 transitions to 20 transitions (Figure 1(a) and Figure 4(a)). As every agent pays a small amount for the good of public punishment, the reward of punishers increased rapidly. Meanwhile, the effect

Figure 3: Basic geographically constrained game

- (a) The evolution of agents: basic geographically constrained game (b) The evolution of rewards: basic geographically constrained game

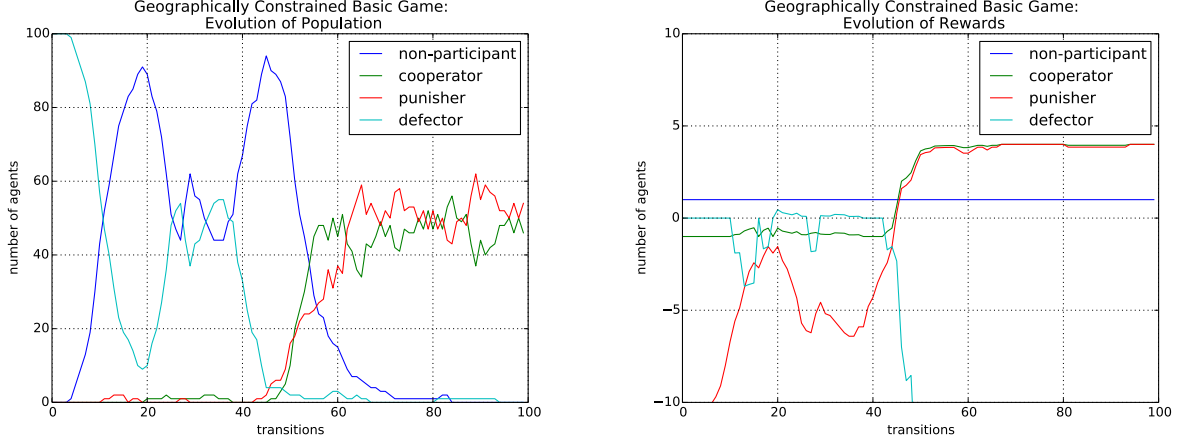
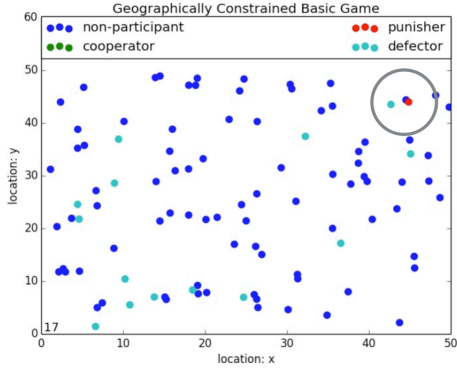
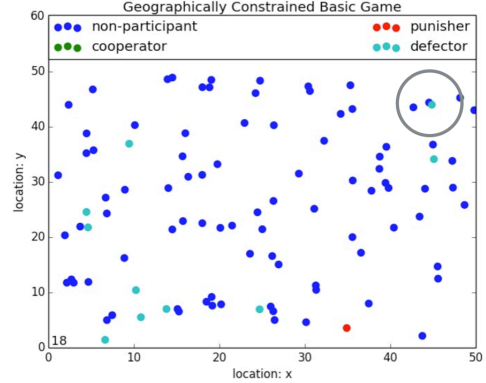


Figure 4: The distribution of agents in geographically constrained game

- (a) The distribution at 17th transition



- (b) The distribution at 18th transition



can be seen in the less fluctuating transient stage at the beginning.

3.4 Effect of geographic constraint and insurance

We have combined geographic constraint and insurance policy in our game model. The impact when both factors work together is largely depend on the parameters, i.e. the radius of geographic constraint and the amount of insurance. These two factors are counteractive, hence the game became more unpredictable. Figure 5 shows one instance of geographically constrained insurance game. Interestingly, there was a plateau period from the start to the 20th iteration in the game initialized by 100 defectors. This phenomenon was caused by the incomplete information. Every agent remained to be defector as there were no more agent seen who better performed. The population started to shift as there were non-participants emerged due to the mutation rate. From

Figure 5: Basic insurance game

- (a) The evolution of agents: basic insurance game (b) The evolution of rewards: basic insurance game

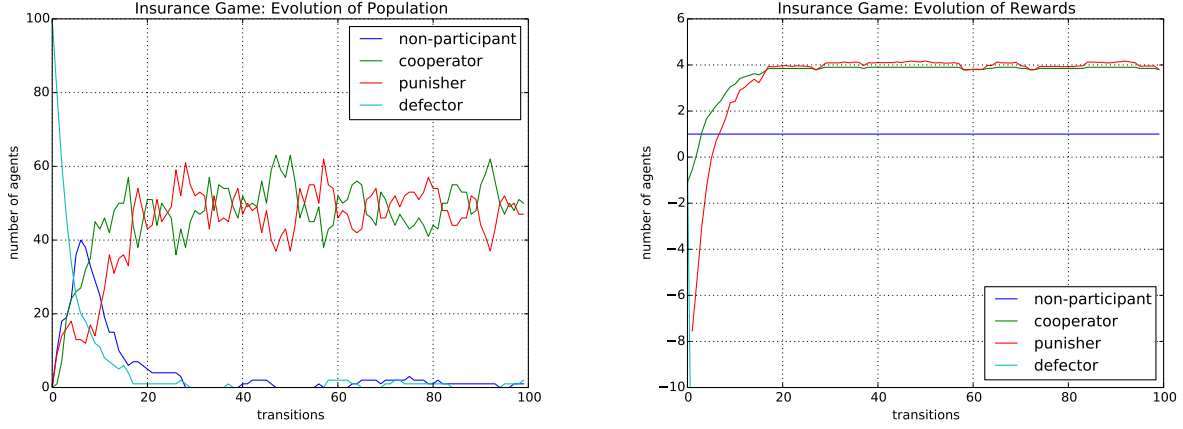
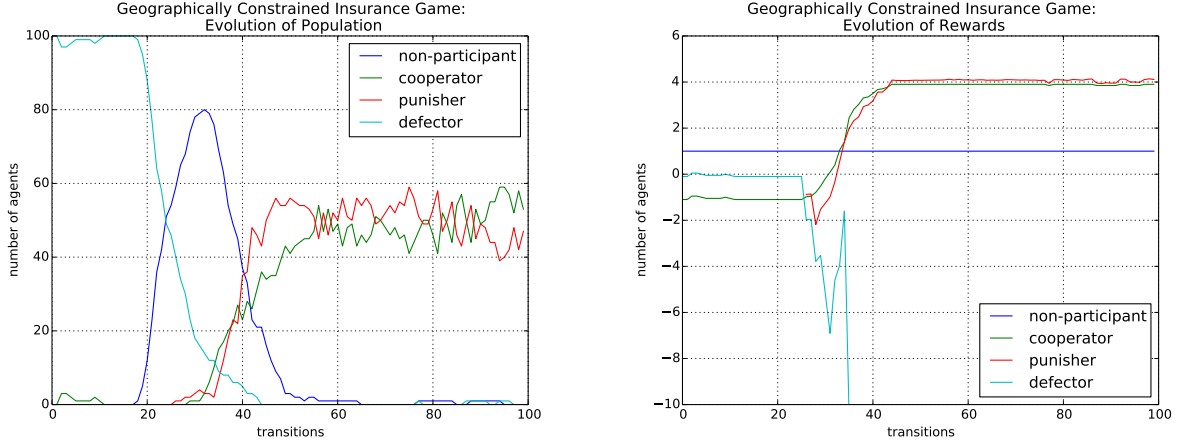


Figure 6: Geographically constrained insurance game

- (a) The evolution of agents: geographically constrained insurance game (b) The evolution of rewards: geographically constrained insurance game



the start to 15th transition, there was a small wave of cooperators. These group of cooperators did not survive since cooperators must depend on punishers to increase their rewards. Insurance policy plays a role in this game such that the punishers evolve faster than in basic geographically constrained games (Figure 2(a) and Figure 5(a)).

4 Discussion

4.1 Different initialization condition

Figure 6 compares two instances of basic geographically constrained insurance game with four different initializations. The punishers always thrived regardless of the initial stages. The only situation where punisher might not be a good choice is when the cooperation is taken over by defectors and non-participation is obviously an alternative. This makes sense since, if a cooperation is so hopeless, it's unreasonable to stay. Unlike basic game, punishers are much stronger and easy to survive in this case.

4.2 How close are models to realities?

The human society is a very complicated system, with far more variables and interactions than can be modeled. It is interesting when we try to recover society behaviors with simple models. Can we really gain any insight or prove anything about reality with these simple models? It would also be interesting to see if we can collect real data, fit the parameters in these simple models. But the question remains, how much we can trust the prediction by these models. Another question to ask is, is there really such an abstraction of one aspect of the society, i.e., can we find a projection onto some low dimensional space of the human society, such that the projection is simple and predictable by a model. The project stimulated us to think a lot about the effectiveness of modeling.

4.3 Sensitivity to parameters

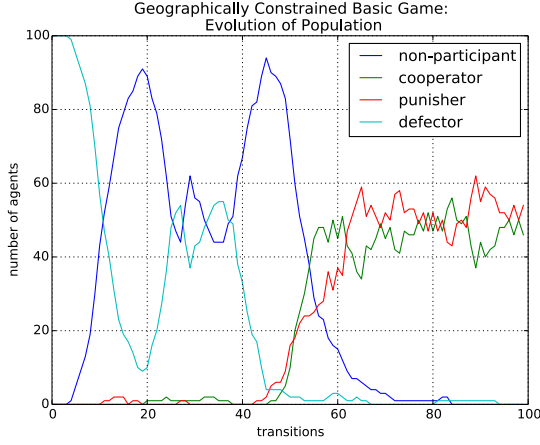
The behavior of insurance games are less sensitive than basic games. However, the parameters have to be within a certain range such that the criterions at the end of Section 2.1 is met. If case of one criterion is broken, such as the 2) one, non-participants will take over and it will be a whole different senario. People might ask that, why is the human society falls into the right region, where punishers and cooperation is the final steady-state? The answer is, we don't. In the history, there emerged many cooperations and punishers whose parameters did not fall in this right region, and they vanished.

5 Conclusion and Future Works

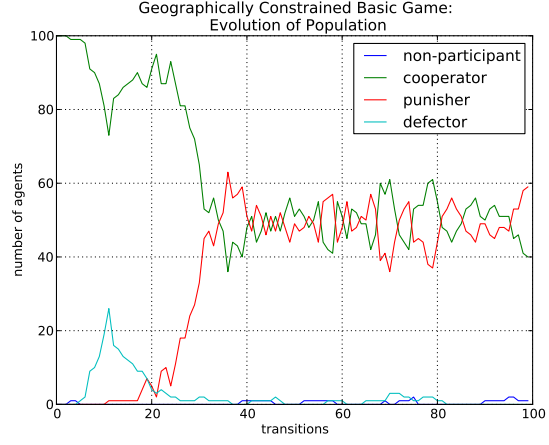
The previous study has demonstrated the mechanism of the thriving of punishers, and pointed out the importance of voluntary participation. We were inspired by this model and would like to incorporate more features in order to explain the two real-world examples: the restricted communication and the enforced tax (or insurance) policy. We have successfully established a model in python and added these new variations. To sum up, our model reiterated the evolutionary advantage of voluntary punishers. We also proved the inhibiting effect of geographic constrain and the catalyzing effect of insurance policy. The model could be utilized in organizational and developmental research such as the optimization of tax policy and the communication strategy in underdeveloped rural area.

Figure 7: Comparison of the initialization

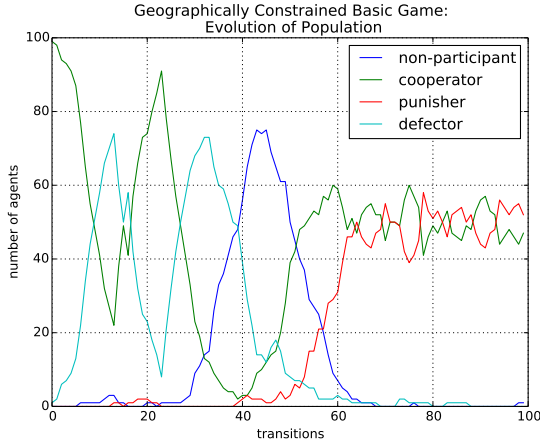
(a) The evolution of agents: basic geographically constrained game (initialized by 100 defectors)



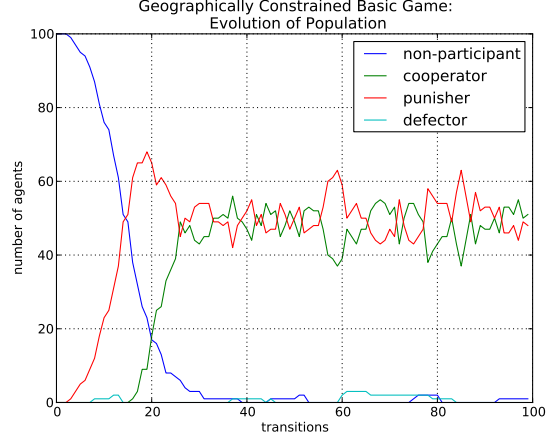
(b) The evolution of rewards: basic geographically constrained game (initialized by 100 cooperators)



(c) The evolution of rewards: basic geographically constrained game (initialized by 99 cooperators and 1 defector)



(d) The evolution of rewards: basic geographically constrained game (initialized by 100 non-participants)



Likewise, our model was based on a series of assumptions. Our model bears the same assumptions with the previous study including excludable cooperative goods, constant pay-off of the non-participants, and ignorance of the economics of the scale. Additionally we have not considered the mobility of the agents, the preservability of public goods, as well as the competition of multiple enterprises. These limitations will be addressed in future works.

5.1 Multiple cooperations

Further on we can split agents according to location or connection constraints into different cooperations, where agents can choose to join different cooperations or stay as non-participant. Cooperations reward can also differ due to their sizes. Incorporating these features, we believe we can simulate interactions between companies, such as large companies gradually expand, turning

everyone into its member and distroying smaller companies. We are also expecting cooperations with punishers are more desirable.

5.2 Accommodation and mobility of agents

In real world, people are always traveling or relocating, changing the dynamics of local world. If we allocate natural resources growth over the whole region, make agents reward also depending on the resource and allow agents to travel, then it would be another interesting simulation.

References

- [1] N. Morris and D. J. Rothman, *The Oxford history of the prison: The practice of punishment in Western society*. Oxford University Press, 1995.
- [2] C. Hauert, A. Traulsen, H. Brandt, M. A. Nowak, and K. Sigmund, “Via freedom to coercion: the emergence of costly punishment,” *science*, vol. 316, no. 5833, pp. 1905–1907, 2007.