Simulation and Modelling of Students Group Dynamics | Coursework 2

Eyad Al-Khayat
University of Southampton
eak1u19@soton.ac.uk

ABSTRACT

The report examines two different solutions for solving a modelling problem related to how students select and change strategies over a period of time and then discuss an extension to the problem. The first solution explores modelling the changes in student dynamics by differential equations. The second approach investigates the use of agent-based modelling. Results from both approaches are discussed and compared.

KEYWORDS

Imitation Dynamics, Replicator Dynamics, Groups Matching Rules

1 INTRODUCTION AND PROBLEM DESCRIPTION

Evolutionary game theory has played an important aspect in biology as it used game theoretic models to see how populations of biological creatures evolved by adapting and following different strategies. What makes Evolutionary game theory different than the classical game theory is that strategies evolve, and secondly it tries to check what are the strategies that are most efficient [3]. The problem this report will explore is considered as an evolutionary game theory problem where we have an initial population of students trying to do projects in groups of n students. At the beginning of each course, individuals choose between two strategies, "H" hard worker, and "L" lazy worker. Students mark depends on the total group effort they have been at and is calculates as follows:

$$e = h \times H + l \times L \tag{1}$$

where:

e = Total Group Effort
h = Number of hard workers in the group
l = Number of lazy workers in the group
H = Effort put by being a hard worker
L = Effort put by being a lazy worker

All students in the group get the same mark by dividing the total group effort by the n number of students in the group.

$$m = e/n \tag{2}$$

Students change their strategies for the next semester by comparing a measure π shown in Equation 3, with another random student. If the performance measure of the selected student is higher then the student's current measure, he/she imitates the strategy of that selected student.

$$\pi = m - a \times S \tag{3}$$

where:

m = Student's mark
a = Parameter chosen at the beginning
S = S=H or S=L based on the strategy

The main questions to be answered for this problem are:

- What is the composition of the population after 4 years?
- What is the composition of the population in the long run?
- How quickly is the equilibrium state reached?
- How do the following parameters influence results:
  - Composition of the population
  - Group size (n)
  - Cost of effort (a)
  - Contribution of workers types to group effort (H, and L)

The rest of the report is structured as follows, Section 2 will show how we can model this problem using differential equations and explores how it can be solved analytically and numerically. Section 3 describes the setup of an agent-based model that mimics the rules and interactions illustrated in the problem description. Section 4 introduces a modification to the problem and explores how this modification will affect the solutions. Section 5, discusses the findings of the simulated models and the challenges faced implementing them.

2 DIFFERENTIAL EQUATIONS MODELLING

Let $M = \{H, L\}$, to be the set of strategies that students can select from each course. We define x_i where $i \in M$ to be the proportion of students following strategy i. We can say that the rate of change of population x_i with respect to time t equals the outflow of students converting from strategy x_i plus the inflow of students converting from the other strategy to strategy x_i. So to model this problem we need two differential equations, one for each strategy. The first simple equation to describe this is:

$$\frac{\partial x_i}{\partial t} = -Inflow + Outflow \tag{4}$$

The behaviour of Equation 4 is somewhat identical to the general replicator equation shown in 5 for an evolutionary population game [1] where the payoff that a student can get will have an impact on how the population will be like in the next course and this is enforced by the replicator dynamics based only on average payoff.

$$\frac{\partial x_i}{\partial t} = x_i(\pi_i - \overline{\pi}) \tag{5}$$

where:

π_i = Payoff of strategy i
$\overline{\pi}$ = Average payoff of the population

The important difference here is that payoffs varies by time according to groups composition. This next section will study this and checks how we can model this difference to the general replicator equation.

2.1 Groups Formation

As described before, at the beginning of each course, individuals are randomly assigned into groups of size n. The composition of students in the groups determine the payoff that each student can get and according to the replicator equation the average fitness of in the population will determine how the proportion of students change in the next course. The process that is followed here to derive the varying group fitness over time is inspired by the work of [2].

A group type is defined as the set of n students in a group. Where g_i is an individual in group g following strategy i, that is either H, or L. The number of ways that student types can form in a group type of size n and two strategies is defined as $\gamma = \frac{(n+2-1)!}{n!}$. For example, for $n = 2$ we have a γ value of 3. This means that we have 3 types of groups. "H,H", "L,L", and "H,L". The proportion of students of type i in group type g is defines as $P_{i_g} = \frac{g_i}{n}$. This could also be described in the matrix $P \in \mathbb{R}^{m \times \gamma}$. So for $n = 2$ and $\gamma = 3$ we get the following matrix describing the proportions of strategies the students are following in each group type:

$$\begin{bmatrix} P_{H\,(H,H)} & P_{H\,(L,L)} & P_{H\,(H,L)} \\ P_{L\,(H,H)} & P_{L\,(L,L)} & P_{L\,(H,L)} \end{bmatrix} = \begin{bmatrix} 2/2 & 0/2 & 1/2 \\ 0/2 & 2/2 & 1/2 \end{bmatrix} \tag{6}$$

Population state is defined as the proportion of strategies at a certain point of time where $x \in \triangle_m$, $\triangle_m = \{x_H, x_L\}$ where X_H, and X_L are the proportions of students following strategy H, and L respectively. Another important term to define is group state which is defined as $z \in \triangle_\gamma$ where z_g is the proportion of groups of type g.

To link the population state with group state at a certain time, a matching rule function is defined, $f : \triangle_m \rightarrow \triangle_\gamma$ such that:

$$x_i = \sum_{g \in G} P_{ig} f_g(x) \tag{7}$$

Finally, $w_{ig}(x)$ denotes the probability that a student of type i is into group of type g:

$$w_{ig}(x) = \frac{P_{ig} f_g(x)}{x_i}$$

2.2 Groups Payoffs

The payoff that students receive by being in a specific group is described by the matrix $A \in \mathbb{R}^{n \times \gamma}$ where A_{ig} is the payoff for a student following strategy i in group type g. For example:

$$A_{H\,(H,H)} = m - \frac{1}{2} = \frac{e}{n} - \frac{1}{2} = \frac{h}{n} - \frac{1}{2} = \frac{2}{2} - \frac{1}{2} = \frac{1}{2}$$

To calculate the total payoff to a student following strategy i in a group type g we can multiply the probability of that student being in that group and we will have:

$$F_i(x) = \sum_{g \in G} w_{ig}(x) A_{ig} \tag{8}$$

Since students are put into groups randomly, [2] describes the uniformly random matching rule for $n = 2$ as follows:

$$f_{(2,0)}(X_H, X_L) = X_h^2 \tag{9}$$

$$f_{(1,1)}(X_H, X_L) = 2X_h X_L \tag{10}$$

$$f_{(0,2)}(X_H, X_L) = X_L^2 \tag{11}$$

2.3 Model Dynamics

Since the composition of groups and students payoffs change by time, this needs to be incorporated into some of the equations identified in previous steps. x^t is defined as the population state at time t which in turn is mapped to group types distribution with the matching rules function $f(x^t$. Under these assumptions Equation 8 becomes as :

$$F_i^{A,f}(x^t) = \sum_{g \in G} w_{ig}(x) A_{ig} \tag{12}$$

The average payoff for both strategies would is defines as follows:

$$\overline{F}^{A,f}(x^t) = \sum_{i=1}^{2} x_i^t F_i^{A,f}(x^t) \tag{13}$$

Plugging the payoffs of groups into the replicator Equation 5 will yield to the following two equations that describe that rate of change of hard and lazy workers populations with respect to time:

$$\frac{\partial x_H}{\partial t} = x_H^t (F_H^{A,f}(x^t) - \overline{F}^{A,f}(x^t)) \tag{14}$$

$$\frac{\partial x_L}{\partial t} = x_L^t (F_L^{A,f}(x^t) - \overline{F}^{A,f}(x^t)) \tag{15}$$

These equations work for any value of group size n, however, the matching function f needs to modified to accommodate the change in group size.

3 AGENT-BASED MODELLING

Modelling the problem as an agent-based approach was much simpler. It was implemented from scratch using Python programming language. Two classes were needed to model the simulation of this problem. The first class (Agent) is related to the student and includes information such as the agent type (hard or lazy worker), the mark of the student, and the performance. Several methods had to deal with switching strategies, and setting and comparing the performance of the student. The second class (Group) kept a record of agents being grouped in the same group. The methods here had to deal with calculating the group effort and then setting the respective performance measure for each agent. Each simulation run involved the implementation of the following steps, at first agents were randomly put into groups based on the group size n. Then, effort for each group was calculated and each student had their mark, and performance. After that, agents had a random pairwise interaction with other agent where they compared their performance measures and only switched strategy if they had a lower value than the other. Initially, the population was split in half. 50% were hard workers and the other half were lazy workers.

The results obtained by this approach followed the same behaviour of the replicator equation where a certain population will grow when a strategy payoff exceeds the average payoff of the population, and decrease in the opposite case. This was true for all values of group size $n > 2$ and with the initial conditions given in the original problem, lazy workers students dominated the population after a relatively small number of runs (around 40). This behaviour is shown in Figure 2. The proportion of the population in the 4th run was 74% of lazy workers, and 26% of hard workers.

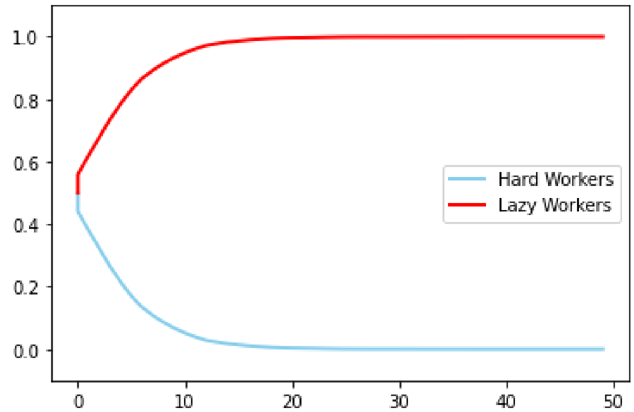

Figure 1: Hard workers dominating the population around run number 40

Group size $n = 2$ had a rather different unexpected behaviour because the proportions of population were fluctuating for a long period, but when running it for more than 25,000 times it finally reached to the same output as before where lazy workers were dominating the population. I assume the reason behind this is that groups of 2 having 2 lazy workers would have a measure π of zero and if the population was composed such group formations lazy workers will always have a bad measure when compared to other hard workers. This would only change when a certain number of group types containing hard and lazy workers are the ones used to map the population, hence making the lazy workers measure above the value of zero.

Figure 2: Fluctuating behaviour of groups with n=2

The cost of effort (a) with group size $n > 2$ had no impact when the value was increasing towards 1. However, when the value is below a certain threshold, that is ($a < 0.24$), the proportions are flipped and hard workers dominated the population. As for Modifying the effort of being a hard or lazy worker (H,L), as long as the value of L is less than H, with group size $n > 2$, lazy workers always ended up taking the whole population. When the value $L > H$ the proportions are flipped and hard workers take over the population. When H and L have equal cost, the population will keep the initial population it started with and will not change at all.

4 PROBLEM EXTENSION

As noticed in the Agent-based modelling solution, lazy workers will dominate the population when the simulation is run for long enough time. This mean that students will never get to pass any course in the future. The modification proposed here is that students will get to remember the mark they had in the previous run and if it was not a passing mark (less than 0.5) the student will switch their strategy with a 50% probability. As with the previous problem, group size of 2 had a different behaviour than values bigger than 2. For $n > 2$, this extension to the problem led to having more balanced proportion of hard and lazy workers over a long time with a tendency of having more hard workers at the end as shown in Figure 3.

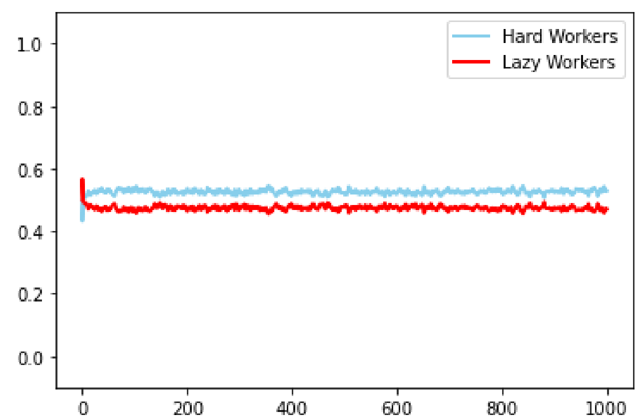

Figure 3: Proportion of hard and lazy workers with n=4

Whereas Figure 4 shows how lazy workers population goes extinct when the group size is $n = 2$.

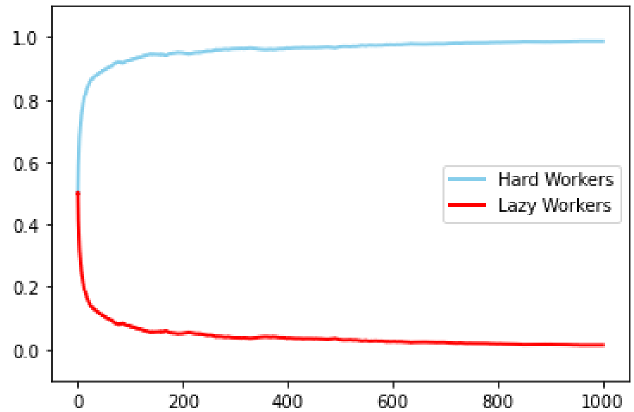

Figure 4: Proportion of hard and lazy workers with n=2

Both of these simulations ran for 1000 times with a population of 10,000 students with the same initial conditions of the original problem. The extension was applied to the agent-based model only. It was quite unclear how this new behaviour could be modeled in the differential Equations identified in 14,and 15. However, implementing this was pretty straight forward in the agent-based approach, as the state of the previous run can be easily saved and checked when forming new groups in the next run. This new extension is closer to the real behaviour of how students deal with the decision of working hard or slacking behind as most of the students will base this decision over their previous marks and hope that they will not repeat the old mistakes by switching to hard worker strategy and securing a better mark.

5 CONCLUSION

In summary, it was noticeable that the agent-based approach was much simpler to implement and model than figuring out the differential equations for the model. It is assumed that both approaches yields similar results, especially when the simulation runs for a long enough time. Modifying and extending the agent-based method is simpler and provides the ability to think of more conditions that will lead to model real world behaviors in a better way.

REFERENCES

[1] Josef Hofbauer and Karl Sigmund. 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press. https://doi.org/10.1017/CBO9781139173179

[2] Martin Kaae Jensen and Alexandros Rigos. 2018. Evolutionary games and matching rules. *International Journal of Game Theory* (2018). https://doi.org/10.1007/s00182-018-0630-1

[3] Carlos P. Roca, José A. Cuesta, and Angel Sánchez. 2009. Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. *Physics of Life Reviews* 6, 4 (2009), 208 – 249. https://doi.org/10.1016/j.plrev.2009.08.001