Replicator Dynamics of an Assignment Game

Ana L. C. Bazzan

PPGC / Instituto de Informática – Fed. Univ. of Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil
`bazzan@inf.ufrgs.br`

Abstract. In games with a high number of players, the main interest is normally on the investigation about the changing in the distribution of individuals along time. In this paper we use this kind of game to model three populations of drivers in a traffic scenario where they have to select routes. This way we investigate the dynamics of the assignment for a non trivial case in which the payoffs of the populations are not symmetric. We show how the convergence to one of the Nash equilibria of the resulting game is achieved when the three populations co-evolve, under different rates of mutants in these populations.

1 Introduction

Evolutionary game theory (EGT) is well-known for investigating the relationship between individual and aggregate behaviors. There are many systems where we nowadays observe a tendency of a complex coupled decision-making process, usually known as collective intelligence. Already in 1950, J. Nash [6] saw this phenomenon, which he then called “mass-action interpretation”. Later, this focus on equilibria was criticized by J. Maynard Smith: ”An obvious weakness of the game-theoretic approach to evolution is that it places great emphasis on equilibrium states ...” [7]. Besides, J. Maynard Smith also dealt with the shift from individual to population level. Even if he borrowed some definitions from standard GT when he introduced the concept of evolutionary stable strategy (ESS) as a way to understand conflicts among animals, he had already noticed that ”there are many situations ... in which an individual is, in effect, competing not against an individual opponent but against the population as a whole... Such cases can be described as ’playing the field’... Such contests against the field are probably more widespread and important than pairwise contests.” [7].

Currently, this kind of modeling is called a population game, which models simultaneous interactions of a large number of simple agents distributed in a finite number of populations. Simple agents here mean that each has a (typically small) number of strategies to choose, causing a minor impact in other agents payoff. Despite this, the payoff of each agent is, as in the classical game theory, conditioned by the distribution of strategies in each population.

In population games, typically, one is not interested in constancy or equilibrium only. Rather, the major interest is on changes. Thus the dynamics of games has been usually the main concerns of EGT. It considers a population of decision-makers (animals, human players) and investigates the rate of change of the strategy profiles in time, as a response to the decisions made by all individuals in the population. The biological interpretation is that a population consists of animals that are genetically programmed to play some strategies. Initially, this strategies may be distributed randomly over the population. The payoff to an individual using a given strategy determines how many offsprings each individual will have in the next generation. This way the composition of individuals (and hence of strategies) in the next generations will change.

The present paper is concerned with one of such systems, namely a traffic network. game theory and evolutionary programming are combined to investigate the dynamics of demand in a traffic network in which the route choice of three populations of agents is modeled in a game-theoretic way. These populations play a normal form game (NFG). This may look trivial. However, a look in the literature shows that, mostly, only two populations are considered, each having only two actions to select, i.e., a $2x2$ game. Moreover, frequently, the game is symmetric. This means that the equations of the replicator dynamics (more on Section 3) have at most two variables. Because we have three populations playing a non-symmetric game, and two of them have three actions to select from, the replicator dynamics is both difficult to analyse and represent.

The goal here is not the computation of the exact equilibrium, especially because in the real-world this is probably a useless effort given that this equilibrium will not last long due to the dynamic nature of the environment. Using the biological metaphor of a population of decision-makers, we consider an environment in which new drivers or agents replace existing ones in a way to reproduce the fact that in real-world networks, drivers unfamiliar with the network (e.g., non-commuters) do use it even if one cannot expect that they have full knowledge about the game. This way, the co-evolution of the three populations is investigated regarding the assignment of routes. The basic idea (further developed in Section 3) is that a population of strategies is reproduced, from generation to generation, proportionally to its fitness.

2 Background and Related Work

In population games, because one deals with a large number of agents, the assumption that agents are able to anticipate what others will do is a strong one. In a traffic network, a driver agent would have to have a complete mental account of all the routes that take it from his origin to his destination. Besides, this agent would have to anticipate all delays that could arise in each of those routes, for any of the possible set of choices of other agents. Nevertheless, game-theoretic approaches to traffic simulation are reported in the literature. In [8], a two-level, three-player game is discussed. Two of these players represent two road authorities, while the population of drivers is seen as the third player. Complete information is assumed, which means that all players (including the population of drivers) have to be aware of the movements of others.

A large number of works about the effect of information on route choice uses abstract scenarios based on static assignment. These abstract scenarios are mostly inspired by congestion or minority games. The basic idea is that agents have to decide simultaneously between two routes; those that select the less crowded one receive a higher reward. Agents' repeated decision-making is coupled to some adaptation or learning strategy so that the next choice is adapted to the reward feedback. Based on this, an equilibrium may be reached. Examples of such abstract two route scenarios can be found in [5, 2, 1]. Also, authors in [3] have investigated the use of minority games to achieve a balanced usage of a road network in which decision are made about which link to follow.

3 Methods

A population game can be defined as follows. $\mathcal{P} = \{1, ..., p\}$ is the society of $p \geq 1$ populations of agents where $|p|$ is the number of populations. $\mathcal{S}^p = \{s_1^p, ..., s_m^p\}$ is set of strategies available to agents in population p. The payoff function is $\pi(s_i^p, \mathbf{q}^{-\mathbf{p}})$.

In this description, $|p|$ populations interact. Agents in population p have m^p possible strategies. Let n_i^p be the number of individuals using strategy s_i^p. Then, the fraction of agents using s_i^p is $x_i^p = \frac{n_i^p}{N^p}$, where N^p is the size of population p. $\mathbf{q}^\mathbf{p}$ is the m^p-dimensional vector of the x_i^p, for $i = 1, 2, ..., m^p$. As usual, $\mathbf{q}^{-\mathbf{p}}$ represents the set of $\mathbf{q}^\mathbf{p}$'s when excluding the population p. The set of all $\mathbf{q}^\mathbf{p}$'s is $\mathbf{q}$. Hence, the payoff of an agent of population p using strategy s_i^p while the rest of the populations play the profile $\mathbf{q}^{-\mathbf{p}}$ is $\pi(s_i^p, \mathbf{q}^{-\mathbf{p}})$.

Originally, EGT derives dynamics for large populations of individuals from *symmetric* two-player games. Consider a (large) population of agents that can use a set of pure strategies $\mathcal{S}^p$. A population profile is a vector σ that gives the probability $\sigma(s_i^p)$ with with strategy $s_i^p \in \mathcal{S}^p$ is played in p.

In two-player games, for *symmetric* Nash equilibria, (σ^*, σ^*), the interpretation related to population biology is as follows: In a population where everyone uses σ^* (the optimal profile), the best is to use σ^* as well.

So far the solution concept proposed by the symmetric EGT, which is basically a population game with pairwise contexts, i.e., two random members of a *single* population meet and play the stage game, whose payoff matrix is symmetric. The reasoning behind these games is that members of a population cannot be distinguished, i.e., two meet randomly and each plays one role but these need not to be the same in each context. Thus the symmetry. However, there is no reason to restrict oneself to a symmetric modeling in other scenarios beyond population biology. For instance, in economics, a market can be composed of buyers and sellers and these may have asymmetric payoff functions and/or may have sets of actions whose cardinality is not the same. In asymmetric games, each agent belongs to one class determining the set of legal strategies. Asymmetric games thus typically involve co-evolution. In the route choice game discussed here, asymmetric games correspond to multicommodity flow (not a single origin-destination pair).

Table 1. Payoff matrices for the 3-player traffic game; payoffs are for player 1 / player 2 / player 3 (the 3 Nash equilibria in pure strategies are indicated in boldface).

	G3		
	G2	S2	B2
G1	1/1/4	**5/6/7**	5/1/7
S1	3/4/6	4/6/8	4/1/8
B1	5/5/7	**5/6/8**	4/0/9

	T3		
	G2	S2	B2
G1	4/4/8	7/4/6	7/1/8
S1	4/6/8	5/4/6	5/1/8
B1	**5/7/8**	5/4/6	4/0/8

Before we introduce the particular modeling of asymmetric population game, we discuss the concept of replicator dynamics. Recall that the composition of the population of agents (and hence of strategies) in the next generations changes with time (in this case generations) suggests that we can see these agents as replicators. The dynamics of this replication is unclear when one simply computes the ESS. Moreover, the ESS may not even exist, given that the set of ESSs is a possibly empty subset of the set of Nash equilibria computed for the NFG. In the replicator dynamics, it is assumed that members of each population p are programmed to adopt one pure strategy from a finite set available to the population, $\mathcal{S}^p = \{s_1^p, ..., s_m^p\}$. Suppose that there are $|p|$ populations of agents, one for each player. The replicator dynamics describes the evolution of the proportion of members of each population playing every strategy.

As previously defined, the fraction of agents using s_i^j is $x_i^p = \frac{n_i^p}{N^p}$. The state of population p can be described as a vector $\mathbf{x^p} = (x_i^p, ..., x_m^p)$. We are interested in how the fraction of agents using each strategy changes with time, i.e., the derivative $\dot{x}_i^p$. Because payoffs represent reproductive fitness that is responsible for the number of successors using each strategy, we can write:

$$\dot{x}_i^p = (\pi(s_i^p, \mathbf{x^p}) - \bar{\pi}(\mathbf{x^p})) \times x_i^p \qquad (1)$$

In Eq. 1, $\bar{\pi}(\mathbf{x^p})$ is the average payoff obtained by p: $\bar{\pi}(\mathbf{x^p}) = \sum_{i=1}^{m} x_i^p \pi(s_i^p, \mathbf{x^p})$.

In the three-population game considered in this paper, to avoid confusion we use the term "player" with its classical interpretation, i.e., the decision-makers of the NFG. Because this game is played by randomly matched individuals, one from each population, we call these individuals "agents". Thus player refers to a population of agents.

We are now in position to use the description of the general population game given in the previous section and instantiate it for our particular scenario. Formally: $\mathcal{P} = \{1, 2, 3\}$; the strategies for each population $p \in \mathcal{P}$ are $\mathcal{S}^1 = \{G1, S1, B1\}$, $\mathcal{S}^2 = \{G2, S2, B2\}$, and $\mathcal{S}^3 = \{G3, T3\}$; the payoff function is given in Table 1.

In order to explain how the fitness function was derived, it is necessary to understand how the three populations interact. Let us assume that we have road network Γ, which is used by three populations of agents to go from their

Table 2. Fraction of agents associated with each (pure) strategy (route choice).

str.	fraction	str.	fraction	str.	fraction
G1	x_1	G2	y_1	G3	z
S1	x_2	S2	y_2	T3	$(1-z)$
B1	$(1-x_1-x_2)$	B2	$(1-y_1-y_2)$		

respective origins to a destination. Each agent must select a route and the payoff obtained is a function of the delay on the selected route.

Now, there are some alternative routes available to each agent in each population. These routes or strategies were named after the following reasoning: G means greedy selection. (G is the most preferred because this route yields the highest utility *if not shared with other populations*); S means second preferred alternative; and B means border route (a route that uses the periphery of Γ). Populations $p = 1$ and $p = 2$ have strategies $\{G1, S1, B1\}$ and $\{G2, S2, B2\}$. Population $p = 3$ is slightly different. It has only two strategies: G3 (greedy) and one called T3 (which calls for agents to turn in order to avoid the low utility when everybody wants to use its greedy option). Combining all these sets, there are 18 possible assignments of routes. We recall that the fraction of agents associated with each $s_i^p \in \mathcal{S}^p$ is given by the vector $x_i^p = \frac{n_i^p}{N^p}$. To facilitate the comprehension of the equations of the replicator dynamics (derived from Eq. 1), as well as to reduce the number of variables, we allow a slight modification in this notation, which is then shown in Table 2.

The delay in each route is the sum of delays on each link in the route. In turn, delays in each link are given by a volume-delay function (VDF) or expression of the number of users of the particular link at a particular time. The VDF used in the present paper considers the number of agents using each link. Specifically, it adds 1 unit each time an agent uses a given link. This way, a link has cost 0 if no agent uses it; cost 1 if one agent uses it; and so forth. One particular link L of the network however does not accommodate all agents. Thus if all agents use L at the same time, each receives a penalty of 1 unit.

As an example, assume that all agents want to use their most preferred route, i.e., they all act greedily and select route G1, G2, and G3 if belonging to population $p = 1$, $p = 2$, and $p = 3$ respectively. An agent in $p = 1$ then has the following cost: 3 (3 links that are not shared) + 2 (one link shared with agents in $p = 2$) + 6 (2 links shared with $p = 2$ and $p = 3$) + 1 (penalty on L). Similar costs apply for all 18 combinations of the routes. The maximum cost is incurred by agents in $p = 2$ when the following combination of route choices is made: B1 / B2 / G3. This cost is 13 for $p = 2$. In order to deal with a maximization (of utility or payoff) problem rather than one of cost minimization, costs are transformed in payoffs as follows. For each joint route choice, the resulting cost for each player is subtracted from the maximum cost (in this case, 13), resulting in the correspondent payoff. Payoffs computed this way are given in Table 1. This way, for the just mentioned example, the payoff of agents in $p = 2$ when the joint choice is B1 / B2 / G3 is not the cost of 13 but in fact 13-13=0.

Table 3. Five Nash equilibria for the three-player traffic game.

Profile	x_1	x_2	$(1-x_1-x_2)$	y_1	y_2	$(1-y_1-y_2)$	z	$(1-z)$	payoff
σ_a	1	0	0	0	1	0	1	0	5/6/7
σ_b	0	0	1	0	0	1	1	0	5/6/8
σ_c	0	0	1	1	0	0	0	1	5/7/8
σ_d	0	0	1	$\frac{2}{3}$	$\frac{1}{3}$	0	$\frac{3}{4}$	$\frac{1}{4}$	
σ_e	0.4744	0	0.5256	0.3863	0.6136	0	0.3524	0.6476	

The values in Table 1 represent an arbitrary assignment of utility of the three players involved, based on the topology of Γ as explained. The utility function $u(.)$ that underlies Table 1 is however equivalent to any other $\hat{u}(.)$ if $\hat{u}(.)$ represents identical preferences of the players, and $u(.)$ and $\hat{u}(.)$ differ by a linear transformation of the form $\hat{u}(.) = A \times u(.) + B$, $A > 0$. Of course equivalence here refers to the solution concept, i.e., a qualitative, not quantitative concept. Equivalent game models will make the same prediction or prescription.

For the three-agent game whose payoffs are given in Table 1, there are five Nash equilibria. While the computation of the three equilibria in pure strategies is trivial, the mixed-equilibria are less obvious. All five Nash equilibria appear in Table 3. In this table, columns 2–4 specify $\mathbf{x^1}$ (fraction of agents selecting each strategy s_i^1 in population $p = 1$), columns 5–7 specify $\mathbf{x^2}$ of $p = 2$, and the last two columns specify $\mathbf{x^3}$ of $p = 3$. This means that, for the first equilibrium (profile σ_a), because $x_1 = 1$, $y_2 = 1$, and $z = 1$, all agents in $p = 1$ selects action G1, whereas all agents in $p = 2$ select S2 and all agents in $p = 3$ select G3.

Regarding the mixed strategy profile σ_d, all agents in $p = 1$ select action B1 (because $x_1 = x_2 = 0$), whereas in $p = 2$, $\frac{2}{3}$ of agents select G2 and $\frac{1}{3}$ select S2. In $p = 3$, $\frac{3}{4}$ of agents select G3 and $\frac{1}{4}$ select T3. Profiles σ_b, σ_c, and σ_e can be similarly interpreted. In the classical GT interpretation of equilibrium, σ_a, σ_b, and σ_c would be Nash equilibria in pure strategies, while the other two equilibria would mean that agents randomize between two pure strategies. The EGT interpretation thought is as follows. If we consider that we are dealing with three populations of agents, we can think about the five equilibria in terms of the percentage of individuals in one of the three populations that in fact select one of the actions available. This seems a more reasonable explanation for the concept of mixed strategies, given that, at each time, players in fact only select an action (a route in the network Γ). It must also be noticed that in asymmetric games, all ESS are pure strategies (for a proof see, e.g., [9]). Thus only σ_a, σ_b, and σ_c are candidates for ESS. Besides, clearly, among σ_a, σ_b, and σ_c, the first two are (weakly) Pareto inefficient because σ_c is an outcome that make all agents better off.

As mentioned, in this paper we are not considering how agents in the three populations of agents indeed select an action. Rather, we are interested in the dynamics of co-evolution of the populations as a whole. For this matter, we assume that agents have a way to learn or adapt (or even any other mechanism of action selection) but this is opaque here. What matters is that they somehow

Table 4. Equations of the replicator dynamics for the traffic network

$$
\begin{aligned}
\dot{x_1} &= x_1(-zy_1 - 2z - 4y_1 + 3 + x_1zy_1 + 2zx_1 + 4x_1y_1 - 3x_1 + zx_2 + 2x_2y_1 - x_2 - y_2 + x_1y_2 + x_2y_2)\\
\dot{x_2} &= x_2(-z - 2y_1 + 1 + x_1zy_1 + 2zx_1 + 4x_1y_1 - 3x_1 + zx_2 + 2x_2y_1 - x_2 - y_2 + x_1y_2 + x_2y_2)\\
\dot{y_1} &= y_1(-zx_1 - 2z - 4x_1 - 2x_2 + 7 + x_1zy_1 + 2zy_1 + 4x_1y_1 + 2x_2y_1 - 7y_1 - 2zy_2 - 4y_2 + x_1y_2 + x_2y_2)\\
\dot{y_2} &= y_2(2z + 4 + x_1zy_1 + 2zy_1 + 4x_1y_1 + 2x_2y_1 - 7y_1 - 2zy_2 - 4y_2 - x_1 - x_2 + x_1y_2 + x_2y_2)\\
\dot{z} &= z(-x_1y_1 - 2x_1 - x_2 - 2y_1 + y_2 + 1 + x_1y_2 + x_2y_2 + x_1zy_1 + 2zx_1 + zx_2 + 2zy_1 - zy_2 - z - zx_1y_2 - zx_2y_2)
\end{aligned}
$$

select an action. This does not mean that agents do not adapt to the environment. In fact they do but because we are assuming that the population is large enough, we may just look at what happens at population level.

To reproduce the behavior of new drivers in the network Γ, we use ideas similar to those in [4] and [10], in which newcomers replace some agents who either leave the population or change their strategies. In either case, newcomers are genetically mutated versions of the agents. To create these, we use a mutation rate p_m: with probability p_m an agent in the population p is replaced by its mutated version. The dynamics of the process is then modeled by a genetic algorithm. Each population p is composed by N^p agents. Each is programmed to play a given strategy $s_i^p \in \mathcal{S}^p$. For instance, an agent in $p = 1$ can be genetically coded to play G1. In each generation, each agent plays g games whose payoff is as in Table 1. The sum of the payoffs obtained by playing these g games is then the fitness of the agent. After these g games are played, the populations of agents are reproduced: In each population p, the fittest agents have a higher probability of being selected. Then each individual suffers mutation with probability p_m, which means that its strategy is changed to another one randomly selected. We are then interested in the replicator dynamics of the agents in the three populations. From Eq. 1, and having in mind that the fractions of agents in each population are denoted as in Table 2, the equations of the replicator dynamics for our specific three-population game appear in Table 4.

4 Experiments and Results

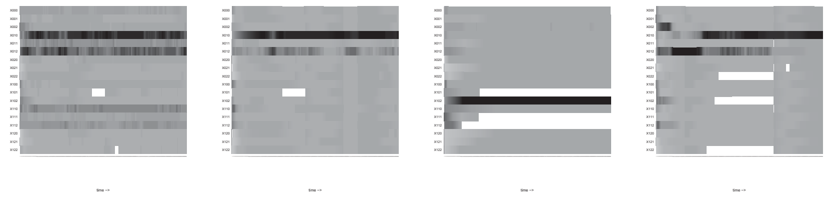
In this section we discuss the numerical simulations of the replicator dynamics for the three-population game described in Section 3. We are interested in investigating issues such as what happens if the populations start with each one using a given profile σ in tames that have more than one equilibrium. For instance, if this profile is σ^*, under which conditions will it remain this way? How many mutants are necessary to shift this pattern? Also, if the population starts using any σ (one of the equilibria), what happens if it is close to (but not actually at) σ^*? Will it tend to evolve towards σ^* or move away? If it reaches σ^*, how long has it taken?

The main parameters of the model, as well as the values that were used in the simulations are: $\mathcal{P} = \{1, 2, 3\}$; $N^1 = N^2 = N^3 = 300$; $g = 10,000$ (number of games played by each agent in each generation); $\Delta = 1000$ (number of generations); and p_m (mutation probability, which varies).

The three Nash equilibria that need to be investigated are, as mentioned, those in pure strategies, i.e., σ_a, σ_b, and σ_c from Table 3. We have analytically checked that only σ_c is an ESS. To do this, it is necessary to analyse which are stable rest points. In simple games, e.g., $2x2$ or even symmetric $2x3$ this can be done graphically. However, because our problem involves five variables (x_1, x_2, y_1, y_2, z), the divergence operator was used. This way, it was verified that the only Nash equilibria where all derivatives are negative is σ_c, meaning that this is a stable rest point. We stress however that such an investigation is only possible for the *static* NFG, i.e., when no change happens in the composition of the populations. When it comes to the dynamics, this investigation can only be done using numerical simulation. For instance, the issue about whether or not the ESS will establish depends on the mutation rate. If it is too high, then the populations never converge to the selection of any Nash equilibria, much less to the ESS, because perturbations happen too often. If the mutation rate is too low, it may be that the initial condition determines which Nash equilibria will establish, which may not be the ESS. In this case, because the mutation is almost zero, the population is not able to shift to the ESS.

Because the replicator dynamics for this problem involves five variables, it is not possible to show the typical replicator dynamics plots that depict the trajectory of the variables, and hence show the convergence to a point. Therefore, as an alternative to show the dynamics of this convergence we use a kind of heatmap, here by means of the software R. In the plots that appear next (which were reduced due to lack of space), heatmaps (here only using shades of gray) are used to convey the idea of the intensity of the selection of each of the 18 joint actions (represented in the vertical axis) along time (horizontal axis). The three Nash equilibria that we seek are represented as `010`, `012`, and `102` for σ_a, σ_b, and σ_c respectively. The darker the shade, the more intense one joint action is selected. Thus we should expect that the three Nash equilibria correspond to the darker strips.

In Figure 1 we show how the selection evolves along time, for high rates of mutation. Figure 1(a) is for $p_m = 10^{-1}$. We can see that althought σ_a (`010`) is clearly more frequently selected, other joint actions also appear often, as, e.g., `012`. Interestingly, their counterparts `110` and `112`, which differ from `010` and `012` by $p = 3$ selecting T3 instead of G3 also appear relatively often. This indicates that agents in $p = 3$ try to adapt to the other two populations. In the end the performance is poor because co-evolution is disturbed by the high rate of mutation (newcomers, experimentation by the agents). This overal picture improves a little with the reduction in p_m. When $p_m = 10^{-2}$ (not shown) and $p_m = 10^{-3}$ (Figure 1(b)), the convergence pattern is clearer but still it is not possible to affirm that one profile has established. We remark that we deliberately show figures where the more frequently selected profile is `010`. However this is not always the case. The other two profiles may also be the most frequently selected. In any case, the general pattern is that there is no clear convergence. When we decrease the rate to $p_m = 10^{-4}$, it is possible to observe that one of the two cases occur: either profile `102` (σ_c) established right in the beginning (Figure 1(c)), or there is a competition between `010` and `012`, with one or the other ending up establishing. In Figure 1(d) it was the case that the joint selection has converged



(a) $p_m = 10^{-1}$ (b) $p_m = 10^{-3}$ (c) $p_m = 10^{-4}$ (d) $p_m = 10^{-4}$

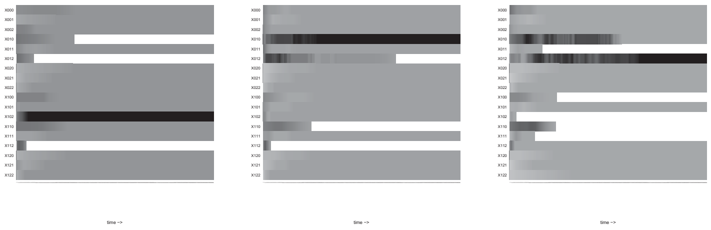
Fig. 1. Evolution of Dynamics for Different p_m.

to `010`, but this needs not to be the case, with `012` being the other convergence possibility. For $p_m = 10^{-5}$ the pattern is pretty much the same as for $p_m = 10^{-4}$. With decrease in p_m, there is a decrease in the time needed to either `010` or `012` establish, if `102` has not already set. For instance, comparing Figure 2(b) to Figure 1(d), one sees that profile σ_a (`010`) established before in the former case. Lower mutation rates follow the same trend.

In short, four conclusions can be drawn from these simulations. First, the co-evolution of the dynamics does not always lead to the ESS computed for the corresponding static NFG. Whereas there is only one ESS among the three Nash equilibria in pure strategies (σ_c), depending on the distribution of strategies in the three populations at the early steps of the game and on the mutation rate, all three Nash equilibria may establish. Second, the mutation rate has an influence on the frequency of selection in the sense that if it is too high, it is not really possible to affirm that one Nash equilibria has established definitively. Third, if the mutation rate is low, then it is possible to see the convergence to one of the three Nash equilibria. Finally, still for low mutation rates, there is a different pattern of convergence when considering the three Nash equilibria. The profile σ_c does establish fast (and does not shift) if it sets at all. When this is not the case, there is a competition between the other two. This competition is determined by agents in $p = 3$: from the payoff matrix (Table 1), one can see that only agents in $p = 3$ have different payoffs in profiles σ_a and σ_b.

5 Conclusion

In this paper, a three-population game was defined in order to model co-evolution of strategies in a scenario in which the payoffs of the populations are not symmetric. Although the game considers three populations only, each having a few actions, we claim that this is not an unrealistic simplification. In fact, in the majority of the situations a traffic engineer has to deal with, there is a small number of commodities (origin-destination pairs) thus three populations is not far from reality. Regarding the number of actions, it is equally the case that in



(a) $p_m = 10^{-6}$ (b) $p_m = 10^{-6}$ (c) $p_m = 10^{-6}$

Fig. 2. Evolution of Dynamics for $p_m = 10^{-6}$.

the majority of the real-world cases drivers do not have more than a handful of options to go from A to B.

One contribution of this paper relies on the modeling, whose analytical solution is not trivial. However the major contribution refers to the aspects related to the dynamics of the co-evolution, as it was shown that the convergence to one of the Nash equilibria is achieved under given mutation rates only. In the future we want to extend this work by combining it with techniques that work at individual agent level as for instance reinforcement learning.

References

[1] A. L. C. Bazzan and F. Klügl. Case studies on the Braess paradox: simulating route recommendation and learning in abstract and microscopic models. *Transportation Research C*, 13(4):299—319, August 2005.

[2] T. Chmura and T. Pitz. An extended reinforcement algorithm for estimation of human behavior in congestion games. *Journal of Artificial Societies and Social Simulation*, 10(2), 2007.

[3] S. M. Galib and I. Moser. Road traffic optimisation using an evolutionary game. In *Proceedings of the 13th annual conference companion on Genetic and evolutionary computation*, GECCO '11, pages 519–526, New York, NY, USA, 2011. ACM.

[4] Kandori, M., G.G. Mailath and R. Rob. Learning, mutation and long run equilibria in games. *Econometrica*, 61:29–56, 1993.

[5] F. Klügl and A. L. C. Bazzan. Route decision behaviour in a commuting scenario. *Journal of Artificial Societies and Social Simulation*, 7(1), 2004.

[6] J. Nash. *Non-Cooperative Games.* PhD thesis, Priceton University, 1950.

[7] J. M. Smith. *Evolution and the Theory of Games.* Cambridge University Press, Cambridge, UK, 1982.

[8] H. J. van Zuylen and H. Taale. Urban networks with ring roads: a two-level, three player game. In *Proc. of the 83rd Annual Meeting of the Transportation Research Board.* TRB, January 2004.

[9] J. N. Webb. *Game Theory – Decisions, Interaction and Evolution.* Springer, London, 2007.

[10] P. H. Young. The evolution of conventions. *Econometrica*, 61:57–84, 1993.