

Social Learning Algorithms Reaching Nash Equilibrium in Symmetric Cournot Games

Mattheos K. Protopapas¹, Francesco Battaglia¹, and Elias B. Kosmatopoulos²

¹ Department of Statistics, University of Rome “La Sapienza”,
Aldo Moro Square 5, 00185 Rome Italy

{matteo.protopapas, francesco.battaglia}@uniroma1.it

² Department of Production Engineering and Management,
Technical University of Crete, Agiou Titou Square
kosmatop@dssl.tuc.gr

Abstract. The series of studies about the convergence or not of the evolutionary strategies of players that use co-evolutionary genetic algorithms in Cournot games has not addressed the issue of individual players’ strategies convergence, but only of the convergence of the aggregate indices (total quantity and price) to the levels that correspond either to the Nash or Walrath Equilibrium. Here we discover that while some algorithms lead to convergence of the aggregates to Nash Equilibrium values, this is not the case for the individual players’ strategies (i.e. no NE is reached). Co-evolutionary programming social learning, as well as a social learning algorithm we introduce here, achieve this goal (in a stochastic sense); this is displayed by statistical tests, as well as “NE stages” evaluation, based on ergodic Markov chains.

1 Introduction

The “Cournot Game” models an oligopoly of two or more firms that decide -independently and simultaneously- the quantities they produce and supply to the market. The total quantity produced by all the firms define -via an exogenous demand function- the equilibrium price at the market. The companies’ goal is to maximize their profit (1), which is the difference of their sales revenue and their production cost. Co-evolutionary Genetic Algorithms have been used for studying Cournot games, since Arifovic [3] studied the cobweb model. In contrast to the classical genetic algorithms used for optimization, the co-evolutionary versions are distinct at the issue of the objective function. In a classical genetic algorithm the objective function for optimization is given before hand, while in the co-evolutionary case, the objective function changes during the course of play as it is based on the choices of the players.

In the algorithms we use here, each chromosome’s fitness is proportional to its profit, as given by

$$\pi(q_i) = Pq_i - c_i(q_i) \tag{1}$$

where $c_i(q_i)$ is the player's cost for producing q_i items of product and P is the market price, as determined by all players' quantity choices, from the inverse demand function

$$P = a - b \sum_{i=1}^n q_i \quad (2)$$

In Arifovic's algorithms [3], populations are updated after every single Cournot game is played, and converge to the Walrasian (competitive) equilibrium and not the Nash equilibrium [2],[15]. Convergence to the competitive equilibrium means that agents' actions -as determined by the algorithm- tend to maximize (1), with price regarded as given, instead of

$$\max_{q_i} \pi(q_i) = P(q_i)q_i - c_i(q_i) \quad (3)$$

that gives the Nash Equilibrium in pure strategies [2]. Later variants of Arifovic's model [5],[7] share the same properties.

Vriend was the first to present a co-evolutionary genetic algorithm in which the equilibrium price and quantity on the market -but not the strategies of the individual players as we will see later- converge to the respective values of the Nash Equilibrium [16]. In his individual learning, multi-population algorithm, which is one of the two algorithms that we study -and transform- in this article, chromosomes' fitness is calculated only after the chromosomes are used in a game, and the population is updated after a given number of games are played with the chromosomes of the current populations. Each player has its own population of chromosomes, from which he picks at random one chromosome to determine its quantity choice at the current round. The fitness of the chromosome, based on the profit acquired from the current game is then calculated, and after a given number of rounds, the population is updated by the usual genetic algorithm operators (crossover and mutation). Since the populations are updated separately, the algorithm is regarded as individual learning. These settings yield Nash Equilibrium values for the total quantity on the market and, consequently, for the price as well, as proven by Vallee and Yildizoglou [15]. In this study, as well as in [11], we have studied the convergence of the individual agents' choices to the Nash Equilibrium quantities.

Finally Alkemade et al. [1] present the first (single population) social learning algorithm that yields Nash Equilibrium values for the total quantity and the price. The four players pick at random one chromosome from a single population, in order to define their quantity for the current round. Then profits are calculated and the fitness value of the active chromosomes is updated, based on the profit of the player who has chosen them. The population is updated by crossover and mutation, after all chromosomes have been used. As Alkemade et al. [1] point out, the algorithm leads the total quantities and the market price to the values corresponding to the NE for these measures.

2 The Models

In all the above models, researchers assume symmetric cost functions (all players have identical cost functions), which implies that the Cournot games studied are symmetric. Additionally, Vriend [16], Alkemade et al. [1] and Arifovic [3] -in one of the models she investigates- use linear (and decreasing) cost functions. If a symmetric Cournot Game, has in addition, indivisibilities (discrete, but closed strategy sets), it is a pseudo-potential game [6] and the following theorem holds:

Theorem 1. *“Consider a n -player Cournot Game. We assume that the inverse demand function P is strictly decreasing and log-concave; the cost function c_i of each firm is strictly increasing and left-continuous; and each firm’s monopoly profit becomes negative for large enough q . The strategy sets S^i , consisting of all possible levels of output producible by firm i , are not required to be convex, but just closed. Under the above assumptions, the Cournot Game has a Nash Equilibrium [in pure strategies]” [6].*

This theorem is relevant when one investigates Cournot Game equilibrium using Genetic Algorithms, because a chromosome can have only a finite number of values and, therefore, it is the discrete version of the Cournot Game that is investigated, in principle. Of course, if one can have a dense enough discretization of the strategy space, so that the NE value of the continuous version of the Cournot Game is included in the chromosomes’ accepted values, it is the case for the NE of the continuous and the discrete version under investigation to coincide.

In all three models we investigate in this paper, the assumptions of the above theorem hold, and hence there is a Nash Equilibrium in pure strategies. We investigate those models for the cases of $n = 4$ and $n = 20$ players.

The first model we use is the linear model used in [1]: The inverse demand is given by

$$P = 256 - Q \tag{4}$$

with $Q = \sum_{i=1}^n q_i$, and the common cost function of the n players is

$$c(q_i) = 56q_i \tag{5}$$

The Nash Equilibrium quantity choice of each of the 4 players is $\hat{q} = 40$ [1]. In the case of 20 players we have, by solving (3), $\hat{q} = 9.5238$. The second model has a polynomial inverse demand function.

$$P = aQ^3 - b \tag{6}$$

and linear symmetric cost function

$$c = xq_i + y \tag{7}$$

If we assume $a < 0$ and $x > 0$ the demand and cost functions will be decreasing and increasing, respectively, and the assumptions of theorem (1) hold. We set $a = -1$, $b = 7.36 \times 10^7 + 10$, $x = y = 10$, so $\hat{q} = 20$ for $n = 20$ and $\hat{q} = 86.9401$ for $n = 4$.

Finally, in the third model, we use a radical inverse demand function

$$P = aQ^{\frac{3}{2}} + b \quad (8)$$

and the linear cost function (7). For $a = -1$, $b = 8300$, $x = 100$ and $y = 10$ theorem (1) holds and $\hat{q} = 19.3749$ for $n = 20$, while $\hat{q} = 82.2143$ for $n = 4$.

3 The Algorithms

We use two multi-population (each player has its own population of chromosomes representing its alternative choices at any round) co-evolutionary genetic algorithms, Vriend's individual learning algorithm [16] and co-evolutionary programming, a similar algorithm that has been used for the game of prisoner's dilemma [10] and, unsuccessfully, for Cournot Duopoly [13]. Since those two algorithms don't, as it will be seen, lead to convergence to the NE in the models under consideration, we introduce two different versions of the algorithms, as well, which are characterized by the use of opponent choices, when the new generation of each player's chromosome population is created, and therefore can be regarded as "socialized" versions of the two algorithms. The difference between the "individual" and the "social" learning versions of the algorithms is that in the former case the population of each player is updated on itself (i.e. only the chromosomes of the specific player's population are taken into account when the new generation is formed), while on the latter, all chromosomes are copied into a common "pool", then the usual genetic operators (crossover and mutation) are used to form the new generation of that aggregate population and finally each chromosome of the generation is copied back to its corresponding player's population. Thus we have "social learning", since the alternative strategic choices of a given player at a specific generation, as given by the chromosomes that comprise its population, are affected by the chromosomes (the ideas should we say) all other players had at the previous generation.

Co-evolutionary programming [13] is quite similar, with the difference that the random match-ups between the chromosomes of the players' population at a given generation are finished when all chromosomes have participated in a game; and then the population is updated, instead of having a parameter (GARate) that defines the generations at which populations update takes place.

In our implementation, we don't use elitism. The reason is that by using only selection proportional to fitness, single (random) point crossover and finally, mutation with fixed mutation rate for each chromosome bit throughout the simulation, we ensure that the algorithms can be classified as *canonical economic GA's* [13], and that their underlying stochastic process form an ergodic Markov Chain [13].

In order to ensure convergence to Nash Equilibrium, we introduce the two "social" versions of the above algorithms. Vriend's multi-population algorithm could be transformed to:

1. A set of strategies [chromosomes representing quantities] is randomly drawn for each player.

2. While *Period* < *T*
 - (a) (If *Period mod GArate* = 0): Use GA procedures (roulette wheel selection, single, random point crossover and mutation), to create a new generation of chromosomes, from a population consisting of the chromosomes belonging to the union of the players' populations. Copy the chromosomes of the new generation to the corresponding player's population, to form a new set of strategies for each player.
 - (b) Each player selects one strategy. The realized profit is calculated (and the fitness of the corresponding chromosomes, is defined, based on that profit).

And social co-evolutionary programming is defined as:

1. Initialize the strategy population of each player
2. Choose one strategy of the population of each player randomly from among the strategies that have not already been assigned profits. Input the strategy information to the tournament. The result of the tournament will decide profit values for these chosen strategies.
3. Repeat step (2) until all strategies are assigned a profit value.
4. Apply the evolutionary operators (selection, crossover, mutation) at the union of players' populations. Copy the chromosomes of the new generation to the corresponding player's population to form the new set of strategies.
5. Repeat steps (2)-(4) until maximum number of generations has been reached.

So the difference between the social and individual learning variants is that chromosomes are first copied in an aggregate population, and the new generation of chromosomes is formed from the chromosomes of this aggregate population. From an economic point of view, this means that the players take into account their opponents choices when they update their set of alternative strategies. So we have a social variant of learning, and since each player has its own population, the algorithms should be classified as "social multi-population economic Genetic Algorithms" [12],[13]. It is important to note that the settings of the game allow the players to observe their opponent choices after every game is played, and take them into account, consequently, when they update their strategy sets.

It is not difficult to show that the stochastic process of all the algorithms presented here form a regular Markov chain [9]. Having a Markov chain implies that the usual performance measures -namely mean value and variance- are not adequate to perform statistical inference, since the observed values in the course of the genetic algorithm are inter-dependent. In a regular Markov chain however, one can estimate the limiting probabilities of the chain by estimating the components of the fixed frequency vector the chain converges to, by

$$\hat{\pi}_i = \frac{N_i}{N} \quad (9)$$

where N_i is the number of observations in which the chain is at state i and N is the total number of observations [4]. In the algorithms presented here, assuming n players, with k chromosomes consisting of l bits in each player's population, the total number of possible states is 2^{knl} , making the estimation of the limiting probabilities of all possible states, practically impossible. On the other hand,

one can estimate the limiting probability of one or more given states, without needing to estimate the limiting probabilities of all the other states. A state of importance could be the state where all chromosomes of all populations represent the Nash Equilibrium quantity (which is the same for all players, since we have a symmetric game). We call this state *Nash State*.

4 Simulation Settings

We use two variants of the three models in our simulations. One about $n = 4$ players and one having $n = 20$ players. We use 20-bits chromosomes for the $n = 4$ players case and 8-bits chromosomes for the $n = 20$ case. A usual mechanism [3],[16] is used to transform chromosome values to quantities. After an arbitrary choice for the maximum quantity, the quantity that corresponds to a given chromosome is given by:

$$q = \frac{1}{q_{max}} \sum_{k=1}^L q_{ijk} 2^{k-1} \quad (10)$$

where L is the length of the chromosome and q_{ijk} is the value of the k_{th} bit of the given chromosome (0 or 1). According to (10) the feasible quantities belong in the interval $[0, q_{max}]$. By setting

$$q_{max} = 3\hat{q} \quad (11)$$

where \hat{q} is the Nash Equilibrium quantity of the corresponding model, we ensure that the Nash Equilibrium of the continuous model is one of the feasible solutions of the discrete model, analyzed by the genetic algorithms, and that the NE of the discrete model will be therefore, the same as the one for the continuous case. And, as it can be easily proven by mathematical induction, that the chromosome corresponding to the Nash Equilibrium quantity, will always be 0101...01, provided that chromosome length is an even number.

The *GArate* parameter needed in the original and the “socialized” versions of Vriend’s algorithms, is set to $GArate = 50$, an efficient value suggested in the literature [15],[16]. We use single - point crossover, with the point at which chromosomes are combined [8] chosen at random. Probability of crossover is always set up to 1, i.e. all the chromosomes of a new generation are products of the crossover operation, between selected parents. The probability of mutating any single bit of a chromosome is fixed throughout any given simulation -something that ensures the homogeneity of the underlying Markov process. The values that have been used (for both cases of $n = 4$ and $n = 20$) are

$$p_m = 0.1, 0.075, \dots, 0.000025, 0.00001.$$

We used populations consisting of

$$pop = 20, 30, 40, 50$$

chromosomes. These choices were made after preliminary tests that evaluated the convergence properties of the algorithms for various population choices, and

they are in accordance to the population sizes used in the literature ([16],[1], etc.).

Finally, the maximum number of generations that a given simulation runs, were

$$T = 10^3, 2 * 10^3, 5 * 10^3, 10^4, 2 * 10^4, 5 * 10^4$$

Note that the number of total iterations (number of games played) of Vriend's individual and social algorithms is $GArate$ times the number of generations, while in the co-evolutionary programming algorithms is number of generations times the number of chromosomes in a population, which is the number of match-ups.

We run 300 independent simulations for each set of settings for all the algorithms, so that the test statistics and the expected time to reach the Nash Equilibrium (NE state, or first game with NE played), are estimated effectively.

5 Synopsis of Results

Although the individual - learning versions of the two algorithms led the estimated expected value of the average quantity (as given in eq.(12))

$$\bar{Q} = \frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n q_{it} \quad (12)$$

(T = number of iterations, n = number of players), close to the corresponding average quantity of the NE, the strategies of each one of the players converged to different quantities. The trajectory of the average market quantity in Vriend's algorithm

$$Q = \frac{1}{n} \sum_{i=1}^n q_{it} \quad (13)$$

(calculated in (13) is quite similar to the trajectory of the same measure in the co-evolutionary case. The estimated average values of the two measures (eq.(12)) were 86.2807 and 88.5472 respectively, while the NE quantity in the polynomial model (6) is 86.9401. The unbiased estimators for the standard deviations of the Q (eq.(14)) were 3.9776 and 2.6838, respectively.

$$s_Q = \frac{1}{T-1} \sum_{i=1}^T (Q_i - \bar{Q})^2 \quad (14)$$

The estimators of the mean values of each player's quantities, as calculated by eq.(15),

$$\bar{q}_i = \frac{1}{T} \sum_{i=1}^T q_i \quad (15)$$

are given on table 1.

That significant difference between the mean values of players' quantities was observed in all simulations of the individual - learning algorithms, in all models

Table 1. Mean Players' Quantities for $n = 4$ players, $pop = 50$, $GArate = 50$, $p_{cr} = 1$, $p_{mut} = 0.01$, $T = 2,000$ generations

Player	Vriend's algorithm	Co-evol. programming
1	91.8309	77.6752
2	65.3700	97.8773
3	93.9287	93.9287
4	93.9933	93.9933

and in both $n = 4$ and $n = 20$, for all the parameter sets used (which were described in the previous section). We used a sample of 300 simulation runs for each parameter set and model, for hypothesis testing. The hypothesis $H_0 : \bar{Q} = q_{Nash}$ was accepted for $\alpha = .05$ in all cases. On the other hand, the hypotheses $H_0 : q_i = q_{Nash}$, were rejected for all players in all models, when the probability of rejection the hypothesis, under the assumption it is correct, was $\alpha = .05$. There was not a single Nash Equilibrium game played, in any of the simulations of the two individual - learning algorithms.

In the social - learning versions of the two algorithms, both the hypotheses $H_0 : \bar{Q} = q_{Nash}$, and $H_0 : q_i = q_{Nash}$ were accepted for $\alpha = .05$, for all models and parameters sets. We used a sample of 300 different simulations for every parameter set, in those cases, as well.

Notice that the all players' quantities have the same mean values (eq. (15)). Mean values of the individual players' quantities on table 2.

Table 2. mean values of the individual players' quantities for $pop = 40$, $p_{cr} = 1$, $p_{mut} = 0.00025$, $T = 10,000$ generations

Player	Social	Social	Individual	Individual
	Vriend's alg.	Co-evol. prog.	Vriend's alg.	Co-evol. prog.
1	86.9991	87.0062	93.7536	97.4890
2	86.9905	87.0089	98.4055	74.9728
3	86.9994	87.0103	89.4122	82.4704
4	87.0046	86.9978	64.6146	90.4242

On the issue of establishing NE in -some- of the games played and reaching the Nash State (all chromosomes of every population equals the chromosome corresponding to the NE quantity) there are two alternative results. For one subset of the parameters set, the social - learning algorithms managed to reach the NE state and in a significant subset of the games played, all players used the NE strategy.

In the cases where mutation probability was too large, the "Nash" chromosomes were altered significantly and therefore the populations couldn't converge to the NE state (within the given iterations). On the other hand, when the mutation probability was low the number of iterations was not enough to have convergence. A larger population, requires more generations to converge to the "NE state" as well. Apparently, the Nash state s_0 has greater than zero frequency in the simulations

that reach it. The estimated time needed to reach Nash State (in generations), to return to it again after departing from it, and the percentage of total games played that were played on NE, are presented on table 3¹.

Table 3. Percentage of the total games in Nash Equilibrium

Model	Algorithm	pop	p_{mut}	T	Gen NE	Ret Time	NE Games
4-Linear	Vriend	30	.001	10,000	3,749.12	3.83	5.54
4-Linear	Co-evol	40	.0005	10,000	2,601.73	6.97	73.82
20-Linear	Vriend	20	.0005	20,000	2,712.45	6.83	88.98
20-Linear	Co-evol	20	.0001	20,000	2,321.32	6.53	85.64
4-poly	Vriend	40	.00025	10,000	2,483.58	3.55	83.70
4-poly	Co-evol	40	.0005	10,000	2,067.72	8.77	60.45
20-poly	Vriend	20	.0005	20,000	2,781.24	9.58	67.60
20-poly	Co-evol	20	.0005	50,000	2,297.72	6.63	83.94
4-radic	Vriend	40	.00075	10,000	2,171.32	4.41	81.73
4-radic	Co-evol	40	.0005	10,000	2,917.92	5.83	73.69
20-radic	Vriend	20	.0005	20,000	2,136.31	7.87	75.34
20-radic	Co-evol	20	.0005	20,000	2,045.81	7.07	79.58

6 Conclusions

We have seen that the original individual - learning versions of the multi - population algorithms do not lead to convergence of the individual players' choices, at the Nash Equilibrium quantity. On the contrary, the "socialized" versions introduced here, accomplish that goal and, for a given set of parameters, establish a very frequent Nash State, making games with NE quite frequent as well, during the course of the simulations. The statistical tests employed, proved that the expected quantities chosen by players converge to the NE in the social - learning versions while that convergence cannot be achieved at the individual - learning versions of the two algorithms. Therefore it can be argued that the learning process is qualitatively better in the case of social learning. The ability of the players to take into consideration their opponents strategies, when they update theirs, and base their new choices at the totality of ideas that were used at the previous period (as in [1]), forces the strategies into consideration to converge to each other and to converge to the NE strategy as well. Of course this option would not be possible, if the profit functions of the individual players were not the same, or, to state that condition in an equivalent way, if there were no symmetry at the cost functions. If the cost functions are symmetric, a player can take note of its opponents realized strategies in the course of play, and use

¹ $GenNE$ = Average number of Generations needed to reach s_0 , starting from populations having all chromosomes equal to the opposite chromosome of the NE chromosome, in the 300 simulations. $RetTime$ = Interarrival Times of s_0 (average number of generations needed to return to s_0) in the 300 simulations. $NEGames$ = Percentage of games played that were NE in the 300 simulations.

them as they are when he updates his ideas, since the effect of these strategies at his individual profit, will be the same. Therefore the inadequate learning process of the individually based learning can be perfected, at the symmetric case. One should note that the convergence to almost identical values displayed in the representative cases of the previous section, holds for any parameter set used in all the models presented in this paper.

References

1. Alkemade, F., La Poutre, H., Amman, H.: On Social Learning and Robust Evolutionary Algorithm Design in the Cournot Oligopoly Game. *Comput. Intell.* 23, 162–175 (2007)
2. Alos-Ferrer, C., Ania, A.: The Evolutionary Stability of Perfectly Competitive Behavior. *Econ. Theor.* 26, 497–516 (2005)
3. Arifovic, J.: Genetic Algorithm Learning and the Cobweb Model. *J. Econ. Dynam. Contr.* 18, 3–28 (1994)
4. Basawa, I.V., Rao, P.: *Statistical Inference for Stochastic Processes*. Academic Press, London (1980)
5. Dawid, H., Kopel, M.: On Economic Applications of the Genetic Algorithm: A Model of the Cobweb Type. *J. Evol. Econ.* 8, 297–315 (1998)
6. Dubey, P., Haimanko, O., Zapechelnyuk, A.: Strategic Complements and Substitutes and Potential Games. *Game Econ. Behav.* 54, 77–94 (2006)
7. Franke, R.: Coevolution and Stable Adjustments in the Cobweb Model. *J. Evol. Econ.* 8, 383–406 (1998)
8. Goldberg, D.E.: *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison - Wesley, Reading (1989)
9. Kemeny, J., Snell, J.: *Finite Markov Chains*. D.Van Nostrand Company Inc., Princeton (1960)
10. Price, T.C.: Using Co-Evolutionary Programming to Simulate Strategic Behavior in Markets. *J. Evol. Econ.* 7, 219–254 (1997)
11. Protopapas, M., Kosmatopoulos, E.: Two genetic algorithms yielding Nash Equilibrium in Symmetric Cournot Games. COMISEF Working Paper Series, WPS-04 (2008)
12. Riechmann, T.: Learning and Behavioral Stability. *J. Evol. Econ.* 9, 225–242 (1999)
13. Riechmann, T.: Genetic Algorithm Learning and Evolutionary Games. *J. Econ. Dynam. Contr.* 25, 1019–1037 (2001)
14. Son, Y.S., Baldick, R.: Hybrid Coevolutionary Programming for Nash Equilibrium Search in Games with Local Optima. *IEEE Trans. Evol. Comput.* 8, 305–315 (2004)
15. Vallee, T., Yildizoglou, M.: Convergence in Finite Cournot Oligopoly with Social and Individual Learning. Working Papers of GRETha, 2007-07 (2007), GRETha, <http://www.gretha.fr> (accessed November 10, 2007)
16. Vriend, N.: An Illustration of the Essential Difference between Individual and Social Learning, and its Consequences for Computational Analyses. *J. Econ. Dynam. Contr.* 24, 1–19 (2000)