Determination of sequential best replies in n-player games by Genetic Algorithms

Mattheos K. Protopapas and Elias B. Kosmatopoulos

Abstract—An iterative algorithm is proposed and tested in Cournot Game models, which is based on the convergence of sequential best responses and the utilization of a genetic algorithm for determining each player's best response to a given strategy profile of its opponents. An extra outer loop is used, to address the problem of finite accuracy, which is inherent in genetic algorithms, since the set of feasible values in such an algorithm is finite. The algorithm is tested in five Cournot models, three of which have convergent best replies sequence, one with divergent sequential best replies and one with "local NE traps"[14], where classical local search algorithms fail to identify the Nash Equilibrium. After a series of simulations, we conclude that the algorithm proposed converges to the Nash Equilibrium, with any level of accuracy needed, in all but the case where the sequential best replies process diverges.

Keywords—Best Response, Cournot oligopoly, Genetic Algorithms, Nash Equilibrium

I. Introduction

We utilize a genetic algorithm to determine a player's best reply in a sequential best reply process context, that converges to a Nash Equilibrium in pure strategies, in Cournot oligopoly games [5]. Lemke and Howson [11] discovered the LCP algorithm that allows for the derivation of the Nash Equilibrium in pure strategies (NE) in a two-player game, but unfortunately, that algorithm cannot be applied in n-player games. Therefore, computational methods are needed in order to calculate the NE's of a n-player game, when an analytical calculation is not possible. Two major categories of this kind of algorithms are evolutionary methods (see for example [13],[16]), which use adaptive agents that represent players, who "learn" using a learning algorithm, and iterative algorithms, which are based on the convergence of the best-reply sequence to a NE, in specific classes of games (see for example [9],[17]).

The algorithm we introduce is an "iterative Nash Equilibrium Search Algorithm" [14]. These algorithms are based on the convergence of the "best reply process" [7] to the Nash Equilibrium in pure strategies (NE). Cournot [5] already proposed an adjustment process for the two-player case, where players decide their quantities sequentially, and each player's decision is the best response to the other player's total output, as it has been formed at the previous round. It's easy to generalize to the n-player case, and the process of the sequential best replies can converge to the NE, under certain provisions

([6], [10], [12], [15]). In the case of convergence to a NE, we then say that the NE is "asymptotically stable" [7].

In the cases in which the best-reply process converges to a NE an iterative algorithm can be used to discover an unknown NE. An optimization algorithm can be used to determine the first player's reply to an initial strategy profile of his opponents. Then that reply is used as the player's strategy in the strategies profile on the next iteration, when the next player's reply will be defined by the optimization algorithm, and so on. That loop ends when the NE has been encountered or a given termination condition holds. These algorithms are known as "iterative NE search algorithms" [14]. The conceptual model of an "iterative NE search algorithm", as given in that article is as follows: First a random strategy is picked as the initial strategy for each player. Then, the profit maximization problem with other players' choices taken as given, is solved for each player and its strategy is updated accordingly. This step is executed for each player in turn, until the NE (as described by an appropriate condition) is found, or another termination condition is met.

Two kinds of problems can arise: the first one arises when the maximum number of iterations is not high enough to determine the NE accurately. A similar cause could be that one of the NE quantities is an irrational number and therefore cannot be discovered with infinite accuracy. This kind of problem cannot be solved in the general case. The only feasible thing to do is to increase the number of iterations, to achieve better accuracy. The second kind of problem is much more serious. As Son and Baldick [14] point out, when a classical local search algorithm -such as Newton's method- is used for the discovery of a player's optimal response (see for example [9]), the best-reply sequence can converge to a "local NE trap" [14], i.e. a strategy profile in which each player's strategy is a local optimum in its optimization problem, but not a global one, and therefore it's not a NE. We use a genetic algorithm as the optimization algorithm in the best-reply determination problem, in order to avoid these "local NE traps". However, since the chromosomes' values belong in a finite set, in any genetic algorithm, an external loop should be utilized to increase the solution's accuracy, and consequently address the problem of the first kind, described earlier, as well.

Mattheos Protopapas is with University of Rome "La Sapienza", Department of Statistics, Aldo Moro Square 5, 00185 Rome Italy, tel. +393391457307, e-mail: matteo.protopapas@uniroma1.it.

Elias Kosmatopoulos is with Technical University of Crete, Department of Production Engineering and Management, Agiou Titou Square, 73100 Chania Greece, tel. +302821037306. e-mail: kosmatop@dssl.tuc.gr.

II. The Algorithm

Since we deal with Cournot oligopoly models [5] the player's strategic choices have to do with quantities of a single, homogeneous product, that are either real or discrete[1]. The

---

[1]In the original Cournot's model, quantities are continuous.

values of the chromosomes of the GA employed can therefore represent real numbers[2]. Since a chromosome's value is an integer (binary) number, a usual decoding scheme is employed [16], adjusted to allow for positive minimum values for the chromosomes.

$$q = q_{min} + \frac{q_{max} - q_{min}}{2^l - 1} \cdot \sum_{n=1}^{l} b_n 2^{n-1} \qquad (1)$$

where b_n is the nth bit of the chromosome, and l is the total number of bits in the chromosome. Adjusted this way, the decoding formula implies that a quantity has a minimum value of q_{min}, a maximum q_{max} and a finite set of values in between, all with equal probabilities.

The algorithm we introduce is as follows:

1) Pick minimum and maximum values for the chromosome values (and consequently, the quantity choice) of each player.
2) Pick n random initial quantities for the n players, creating an initial profile of strategies.
3) For each player, solve its profit maximization problem, using a "canonical Genetic Algorithm" [8].
4) Update the corresponding player's strategy in the profile of active strategies.
5) Loop steps 3-4 for the next player, until all players' strategies have been determined.
6) Loop to step 3, unless the profile of active strategies has remained unaltered (so NE is assumed to be found) or a given number of maximum iterations has been reached.
7) If NE is assumed to be found, terminate.
8) If the accuracy of the current solution is considered to be adequate, terminate. Else update the minimum and maximum values of the players' chromosomes, to allow a better refinement of the solution space and repeat from step 1.

The "canonical GA" is a single-population genetic algorithm with the probability for a chromosome to be selected as a parent being proportional to its fitness (roulette wheel selection), single-point crossover with fixed probability of application, fixed mutation probability throughout the GA process and no elitism, i.e. the next generation is consisted entirely of the offspring of the current chromosome population [8]. We also use "ordered" fitness in our implementation (chromosomes are ranked on the basis of their implied profits and fitness values ranging from 1 to the number of the chromosomes in the population, are attributed to them).

Since the strategy of a player is updated in the active strategies profile each time its profit maximization problem is addressed, we have a "sequential best-reply correspondence" [6], where each player determines its strategy and the next player considers as given that updated strategy, so "responds" in a sense, to that choice. It is possible to adjust the algorithm to function with "simultaneous best replies" [6], or with different opponents' quantities to be considered as given from the active player, when its profit maximization problem is addressed[3]. In any case, the "sequential best reply correspondence" converges to a NE under a broader set of

assumptions in the Cournot case or in any game with strategic substitutes in general [3],[6].

The convergence of the best-replies sequence to the NE, also means that the accuracy of the algorithm is better for a larger number of iterations in the internal loop (steps 3-5), i.e. the values of the quantities in the sequence will be closer to the NE values. Therefore, one can use the minimum and maximum values attained for any players' quantity in the current internal iteration, as the minimum and maximum allowable values for the chromosomes in the outer loop (steps 1-7). In practice a qualitatively assessment of the situation can provide better results with a smaller number of external iterations, and that's the tactic we employ.

III. THE MODELS

A linear Cournot model introduced by Alkemade et al. [1] is used as an initial testbed. The polynomial and exponential models used in Protopapas and Kosmatopoulos [13] are also used, adjusted in the cost functions. The new cost functions are not the same for all players, allowing the study of non-symmetric Cournot games. The inverse demand in the first model is given by

$$P = 256 - Q \qquad (2)$$

with $Q = \sum_{i=1}^{n} q_i$, while cost functions are the same for the n players:

$$c(q_i) = 56q_i \qquad (3)$$

The polynomial Cournot model introduced introduced by the authors [13] has inverse demand function:

$$P = aQ^3 - b \qquad (4)$$

where $a = -1$ and $b = 7.3710^7$ and the same common cost functions, as in the previous model. In the exponential model the inverse demand function is [13]

$$P = aQ^{3/2} - b \qquad (5)$$

a, b being equal to the previous case. In this implementation of the two models we assume different cost function for each player:

$$c_k = kxq_i + y \qquad (6)$$

where k is the index of the respective player, $x = 10$ and $y = 10$, addressing the case of non-symmetric Cournot games. In all three models we study the $n = 4$ player case. We used MATLAB Optimization Toolbox to discover the unique NE in the polynomial and exponential models (table I). For the linear model, NE is derived easily and it is $q^N = 40$ for each one of the 4 players .

TABLE I
NE IN POLYNOMIAL AND EXPONENTIAL MODELS

Player	Polynomial	Exponential
1	86.9400905	86.75703497
2	86.94006293	83.05809161
3	86.94003537	79.35914826
4	86.94000781	75.66020491

In all these models, the assumptions that, as proved in [6], guarantee the convergence of the sequential best reply process

---

[2]If the non-continuous version of the Cournot Game is addressed, a "1-1" correspondence between chromosome values and allowed quantities can be utilized to simplify the algorithm.

[3]Fundenberg and Tirole [7] for example present the case when the expected quantity chosen by any other player is a weighted average of the quantities chosen in the past from that player

to a NE, hold and, furthermore, the NE is unique (as it can be easily proven using elementary calculus), and no local optima ("local NE traps") exist.

A model introduced by Arifovic [2] is used for the $n = 2$ case, with parameters chosen in a way such that a) a NE exists and b) the best-reply sequence diverges from that NE. A sufficient condition for the latter [7] is

$$\frac{\partial^2 u_1}{\partial q_1 \partial q_2} \frac{\partial^2 u_2}{\partial q_1 \partial q_2} > \frac{\partial^2 u_1}{\partial^2 q_1^2} \frac{\partial^2 u_2}{\partial^2 q_2^2} \qquad (7)$$

where u_1, u_2 are the payoff functions for each player, which are functions of the chosen quantities q_1, q_2. The inverse demand function in Arifovic's [2] model are

$$P = A - B(q_1 + q_2)$$

and each player's cost is given by

$$c = xq_i + yq_i^2$$

The payoff functions are, therefore

$$u_i = Pq_i - c_i = [A - B(q_i + q_j)]q_i - xq_i - yq_i^2$$

where $q_i = \{1, 2\}$ and $q_j = \{2, 1\}$. So,

$$\frac{\partial u_i}{\partial q_i} = A - B(q_i + q_j) - Bq_i - x - 2yq_i$$

and

$$\frac{\partial u_i}{q_i} = 0 \Leftrightarrow q_i = \frac{A - x - Bq_j}{2B + 2y}$$

From the latter, we derive the "reaction functions" that determine player i's best response to a given quantity choice by player j,

$$r_i(q_j) = \frac{A - x - Bq_j}{2B + 2y} \qquad (8)$$

and the Nash Equilibrium quantities

$$q_i^N = \frac{A - x}{3B + 2y} \qquad (9)$$

Since a maximum is required, the second order partial derivatives of the payoff functions must be negative at the NE quantities

$$\frac{\partial^2 u_i}{\partial^2 q_i} < 0 \Leftrightarrow B + y > 0$$

and from 7

$$\frac{\partial^2 u_1}{\partial q_1 \partial q_2} \frac{\partial^2 u_2}{\partial q_1 \partial q_2} > \frac{\partial^2 u_1}{\partial^2 q_1^2} \frac{\partial^2 u_2}{\partial^2 q_2^2} \Leftrightarrow B^2 > 4(B + y)^2$$

Since $B + y > 0$ we have

$$B^2 > 4(B + y)^2 \Leftrightarrow y < \frac{-B}{2}$$

and finally,

$$-B < y < \frac{-B}{2} \qquad (10)$$

For these values of y there is a NE and the best replies sequence diverges from it, at the same time. For the values $A = 1000, B = 5, x = 100$ and $y = \frac{-10}{3}$ used for simulations, the best reply sequence diverges (see fig. 1).

Finally, we use one of the "quasi-Cournot" models used in [14]. Although the models employed there are not Cournot
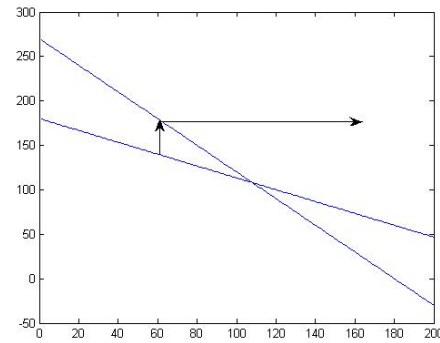

Fig. 1. Best replies in Arifovic's model

Games per se, since the players have different demand functions, existence of "local NE traps" makes them quite interesting to study. As Son and Baldick prove [14], a "local search algorithm" can get stuck to a "local NE trap" instead of finding the absolute NE. The model we use (introduced by Son and Baldick [14]) is a duopoly. The two players' payoff functions are

$$u_1(q_1, q_2) = 21 + q_1 \sin \pi q_1 + q_1 q_2 \sin \pi q_2$$

$$u_2(q_1, q_2) = 21 + q_2 \sin \pi q_2 + q_1 q_2 \sin \pi q_1$$

and, as reported in [14] the NE quantities are[4] $(q_1, q_2) = (6.61, 6.61)$ and the "local NE traps" where the local search algorithms ([9],[17]) could erroneously converge to are $(6.58, 4.66)$, $(0.25, 0.92)$ and $(6.54, 2.8)$.

IV. RESULTS

A. Linear Model

For minimum quantity $Q_{min} = 0$, maximum quantity $Q_{max} = 120$ for all players, population size $pop = 100$ chromosomes, probability for crossover $p_c = 1$ and probability of mutation of a bit $p_m = 0.01$ we got the following results (table II[5]).

TABLE II
LINEAR MODEL

88.554	53.287	40.964	38.243	38.555	39.274
47.109	50.2	45.586	41.915	40.313	39.844
28.269	39.233	42.405	42.16	41.146	40.444
18.047	28.64	35.522	38.841	39.993	40.219
39.747	39.945	40.02	40.026	40.015	40.005
39.795	39.869	39.946	39.987	40.001	40.004
40.12	39.995	39.969	39.977	39.989	39.997
40.195	40.095	40.032	40.005	39.997	39.997
40.001	40	39.999	40	40	40
40.003	40	40	40	40	40
40	40	40	40	40	40
39.998	39.999	40	40	40	40

[4] We discovered, after analyzing the model, that these values are approximations of the respective quantities, to the second decimal digit.

[5] Each line represents the respective player's quantity at the given round. Iterations 1-10 are in the first block and iterations 11-20 in the second

The algorithm converged to the NE, after the 16th iteration. Allowing the algorithm to continue until the 1000th iteration, we had[6] expected frequency of the state defined by the NE strategy (all players choose $Q_i = 40$) $\hat{\pi}_N = 0.984$ and expected return time to this state $\hat{m}_{NN} = 1.01626$.

B. Polynomial Model

Using the same parameter set as in the linear case, we observe the process does not converge to a NE (for 1000 iterations of the inner loop) with an adequate level of accuracy. Chosen quantities (after an "initial" number of 100 iterations have passed) range between 86 and 87.5 (see fig. 2).
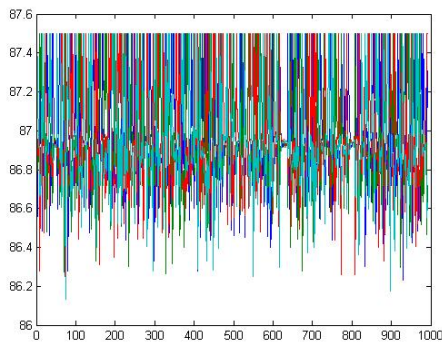

Fig. 2. Polynomial model. First outer loop.

A second outer loop iteration is utilized with $Q_{min} = 86$ and $Q_{max} = 87.5$ for all 4 players. The quantities chosen by each player, for iterations $100, ..., 1000$, range between the values seen on table III:

TABLE III
POLYNOMIAL MODEL. SECOND OUTER LOOP

	player 1	player 2	player 3	player 4
min	86.93750	86.93750	86.93750	86.93750
max	86.94310	86.94400	86.94460	86.94370

Setting the minimum and maximum allowable quantities for the players equal to those values acquired on table III and executing another iteration of the outer loop, we got the maximum and minumum quantities in the iterations $100, ..., 1000$ of the inner loop shown on table IV and fig. 3.

TABLE IV
POLYNOMIAL MODEL. THIRD OUTER LOOP

	min	max
player 1	86.940046425672900	86.940162502539200
player 2	86.940012996781300	86.940162502539200
player 3	86.939973528073800	86.940162502539200
player 4	86.939924654316600	86.940019043368400

We could continue this process of refining the minimum and maximum quantities and executing more iterations of the outer loop of the algorithm, if better accuracy was required.

---

[6]The random process underlying a "canonical" Genetic Algorithm is an ergodic Markov chain, and the same holds, consequently, for the random process defined by the strategies chosen.
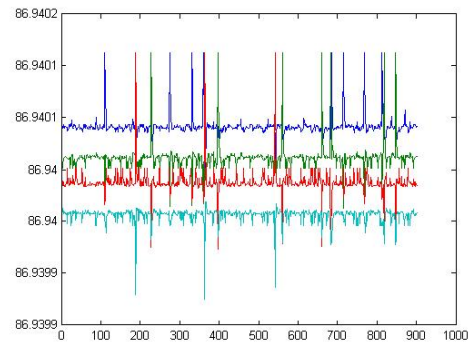

Fig. 3. Polynomial model. Third outer loop.

C. Exponential Model

For $pop = 100$, $p_c = 1$, $p_m = 0.01$, and initial $Q_{min} = 0$, $Q_{max} = 120$ we have the results shown on table V and fig. 4:

TABLE V
EXPONENTIAL MODEL. FIRST OUTER LOOP.

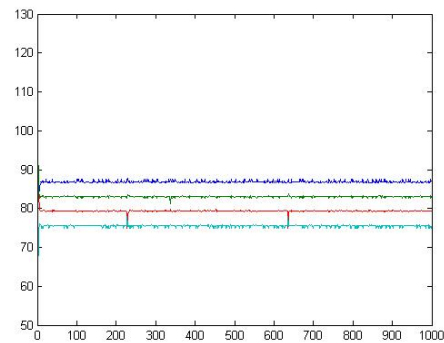|	min	max
player 1	86.620985623346000	87.880599861717100
player 2	81.249886751066900	83.759387740505000
player 3	74.999880790596800	79.931716853825400
player 4	74.999880790596800	77.953794435305100

Fig. 4. Exponential model. First outer loop.

Repeating the outer loop for minumum and maximum allowable values, according to table V we had the results shown on table VI, and fig. 5.

TABLE VI
EXPONENTIAL MODEL. SECOND OUTER LOOP.

	min	max
player 1	86.718436392246600	86.777500150203800
player 2	83.037370488520100	83.066017366425900
player 3	79.321249409913400	79.363343823760800
player 4	75.639685287175500	75.732500708103900

Finally, we executed a third iteration of the outer loop, setting the minimum and maximum allowable values equal to those of table VI. The results are shown on table VII.
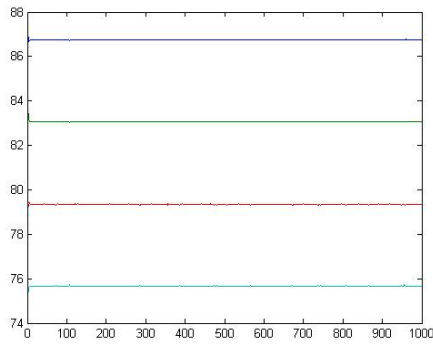
Fig. 5.   Exponential model. Second outer loop.

TABLE VII
EXPONENTIAL MODEL. THIRD OUTER LOOP.

	min	max
player 1	86.755675667069100	86.757312379383400
player 2	83.057409410926800	83.058247337057300
player 3	79.358082017011800	79.359397508843500
player 4	75.659988591350600	75.662889164536500

A comparison between the minimum and maximum values of the players, attained in the third iteration of the outer loop leads to the conclusion (also seen by comparing tables VII and I) that the algorithm gives the NE quantities with an accuracy of 4 decimal digits.

D. Model with divergent best replies

We used various parameter sets for the simulations in the divergent model, inspired by Arifovic [2]. In all cases, the algorithm diverges. In fig. 6 we see the outcome for $Q_{min} = 0$ and $Q_{max} = 300$, for both players.
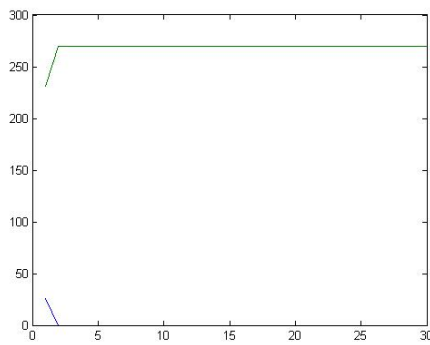

Fig. 6.   Divergent model.

Since the best response to $q_i = 0$ is $q_j = 270$ and to $q_j = 270$ it is $q_i = -135 < 0$ (as derived from eq.8), the situation depicted in fig. 6 is easily understood.

E. Model with local NE traps

We used the parameters $pop = 100$, $p_c = 1$, $p_m = 0.01$, and $Q_{min} = 0$, $Q_{max} = 7.5$ for each of the two players. We

run 100 simulations[7], starting from different quantities in the initial strategy profile, and the algorithm always converged to the NE quantities (within a given degree of accuracy of course). So the local NE traps were avoided in all of the 100 runs of the simulation conducted. On table VIII we see the first 10 iterations of the inner loop for two of those runs.

TABLE VIII
LOCAL NE TRAPS MODEL. INITIAL ITERATIONS FOR 2 RUNS

iter	player 1	player 2	player 1	player 2
1	0	6.514891639	6.581954081	6.613768686
2	6.617437952	6.611243831	6.611615764	6.611830341
3	6.611866104	6.611801731	6.611808883	6.611808883
4	6.611808883	6.621100064	6.611808883	6.611808883
5	6.610857592	6.611901867	6.611944782	6.611794578
6	6.611801731	6.611808883	6.611944782	6.621100064
7	6.611944782	6.611794578	6.610857592	6.611901867
8	6.611944782	6.611801731	6.562499106	6.621100064
9	6.,611944782	6.611944782	6.61085044	6.611901867
10	6.611794578	6.611808883	6.611801731	6.611808883

V. DISCUSSION

The algorithm introduced here is quite effective when the theoretical optimal response sequence converges to the NE. The algorithm converged to the NE[8], in all of the models studied here, in which this precondition holds. The "linear", "polynomial" and "exponential" models are games with "strategic substitutes" [4], and consequently sequential best replies converge to the NE ([3], [6]). Another trivial necessary condition is that the NE quantities belong to the feasible set of the chromosome values. This is not always possible, since the NE quantities might sometimes be irrational or have a large number of decimal digits, while a typical encoding scheme as the one used, implies the chromosomes encode values with finite number of digits. In those cases the outer loop we employed leads to increased accuracy of the computed NE. Those are the cases of the "polynomial" and "exponential" models, where increased accuracy is achieved, after the fine-tuning process of the chromosome values bounds, and the consequent outer loop iterations. Of course, if the Cournot game studied is of the "discrete type" (players can only use discrete quantities), that problem does not exist, provided that the chromosomes have enough digits to hold all the possible admissible quantity values.

The above remarks also hold in the model with "local NE traps". As pointed earlier, the quantities of the NE and the local NE traps are given with a finite accuracy of two decimal digits by [14]. The algorithm attains that level of accuracy relatively easily (with $Q_{min} = 0$, $Q_{max} = 7.5$ less than 10 iterations of the inner loop are generally needed to set correctly the three first digits of the players' quantities). If better accuracy is required more outer loop iterations can be employed.

The only case when the algorithm clearly diverges, is the case of the model with divergent best replies process used. By using $Q_{min} = 0$ and $Q_{max} = 300$[9], the quantities proposed by

[7]Just one iteration of the outer loop. The minimum and maximum quantities were not updated

[8]within a finite level of accuracy

[9]and random initial quantities as always

the algorithm diverge quickly after the initial iterations of the inner loop, something that leads to a situation where one of the players "picks" the zero quantity. Since the best reply to this quantity is 270, which is correctly identified by the algorithm, and the other player assumes that choice, the best reply of his opponent is then an inadmissible negative quantity. The best reply under the constraint $Q_{min} = 0$ is again zero, so the algorithm attributes this value to the other player correctly, thus entering an infinite path of attributing 0 to one of the players and 270 to the other, ad infinitum.

VI. Conclusions

The potency of the genetic algorithm meta-heuristic is evident in the case of Cournot games, as in most other cases genetic algorithms have been applied. The convergence of the sequential best reply search is almost ensured when a genetic algorithm is used for the search of a player's best response to other players strategies, provided that the theoretical best reply sequence converges to a Nash Equilibrium. This depends on the model, so if appropriate conditions hold for the demand and cost functions, such that the sequence of iterative best replies can be proved to an existing Nash equilibrium then the algorithm should derive that strategy profile. One can also study the case of simultaneous best replies, or a case in which the expected opponents' quantities are different of those selected by the players at the previous round.

We have also seen that even if there are "local NE traps", a case in which classical local search algorithms fail, a genetic algorithm will converge to the "true NE". The only case when the GA did not find the Nash Equilibrium profile is the case when the sequence of best replies diverges. In that case, the quantities proposed by the algorithm also diverge. Finally, the drawback of the accuracy of the genetic algorithm, a usual case in continuous problems, is addressed by introducing an outer loop in which the values sets of the chromosomes are adjusted to improve the accuracy of the calculated NE, after each iteration.

REFERENCES

[1] F. Alkemade, H. La Poutre and H. Amman, "On Social Learning and Robust Evolutionary Algorithm Design in the Cournot Oligopoly Game", *Computational Intelligence*, vol. 23, 2007, pp. 162–175.

[2] J. Arifovic,"Genetic Algorithm Learning and the Cobweb Model", *Journal of Economic Dynamics and Control*, vol. 18, 1994, pp. 3–28.

[3] R. Amir, "Cournot oligopoly and the theory of supermodular games", *Games and Economic Behavior*, vol. 15, 1996, pp. 132-148.

[4] J.I. Bulow, J.D. Geanakoplos and P.D. Klemperer, "Multimarket oligopoly: strategic substitutes and complements", *Journal of Political Economy*, vol. 93, 1985, pp. 488-511.

[5] A.A. Cournot, *Recherches sur les principes mathematiques de la theorie des richesses (Researches into the Mathematical Principles of the Theory of Wealth)*, English translation by N.T. Bacon, 1897.

[6] P. Dubey,O. Haimanko and A. Zapechelnyuk, "Strategic Complements and Substitutes and Potential Games", *Games and Economic Behavior*, vol. 54, 2006, pp. 77–94.

[7] D. Fundenberg and J. Tirole, *Game Theory*, The MIT Press, Cambridge MA, 1991.

[8] D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA, 1998.

[9] B.F. Hobbs, C.F. Metzler and J.S. Pang, "Strategic gaming analysis for electric power systems: An MPEC approach", *IEEE Transactions on Power Systems*, vol. 15, 2000, pp. 638-645.

[10] N.S. Kukushkin, "A fixed-point theorem for decreasing mappings", *Economic Letters*, vol. 46, 1994, pp. 23-26.

[11] C.E. Lemke and J.T. Howson, "Equilibrium points of bimatrix games", *Journal of the Society for Industrial and Applied Mathematics*, vol. 12, 1964, pp. 413-423.

[12] P. Milgrom and J. Roberts, "Rationalizability, learning, and equilibrium in games with strategic complementarities", *Econometrica*, vol. 58, 1990, pp. 1255-1277.

[13] M. Protopapas and E. Kosmatopoulos, "Two Genetic Algorithms yielding Nash Equilibrium in Cournot Games", Technical University of Crete Working Paper 29-01-2008.

[14] Y.S. Son and R. Baldick, "Hybrid Coevolutionary Programming for Nash Equilibrium Search in Games with Local Optima", *IEEE Transactions on Evolutionary Computation*, 2004, vol. 8, pp. 305–315.

[15] M. Voorneveld, "Best-response potential games, *Economic Letters*, 2000, vol. 66, pp. 289-295.

[16] N. Vriend, "An Illustration of the Essential Difference between Individual and Social Learning, and its Consequences for Computational Analyses", *Journal of Economic Dynamics and Control*, 2000, vol. 24, pp. 1–19.

[17] J.D. Weber and T.J. Overbye, "A two-level optimization problem for analysis of market bidding strategies",. In *Proceedings of Power Engineering Society Summer Meeting*, 1999, vol. 2, pp. 846–851.