

# Spatial anisotropy encourages parapatric speciation in Artificial Evolution

*SIMONCINI David<sup>1</sup> and DEFOIN-PLATEL Michael<sup>1,2</sup>*

<sup>1</sup>Laboratoire d'Informatique Signaux et Systèmes de Sophia Antipolis  
UNSA-CNRS  
2000, route des lucioles, BP121  
06903 Sophia Antipolis, France

<sup>2</sup>Knowledge Engineering and Discovery Research Institute  
AUT  
Auckland, New Zealand

Presented at SIRC 2006 - The 18<sup>th</sup> Annual Colloquium of the Spatial Information Research Centre  
University of Otago, Dunedin, New Zealand  
November 6<sup>th</sup>-7<sup>th</sup> 2006

## ABSTRACT

The anisotropic selection in Cellular Genetic Algorithms (cGAs) is known to allow the control of the selective pressure but also to enhance the diversity which are two important issues in Genetic Algorithms, especially when trying to solve difficult optimization problems. In cGAs, the population is embedded upon a grid. With the spatial anisotropy of the neighborhood of the each grid cell, new evolutionary dynamics such as speciation can be expected. The process of species formation in the presence of some gene flow between diverging population is called parapatric speciation. In this paper, the artificial evolution of populations in cGAs in the presence of anisotropic selection is compared to this natural speciation phenomenon.

**Keywords and phrases:** Evolutionary algorithms, Structured Population, Speciation

## 1 Introduction

Evolutionary Algorithms (EA) are search methods that handle populations of potential solutions. In EA, the individuals undergo alternated phases of selection and variation. The selection phase is based on the evaluation of the performances of each individual, whereas the variations are performed by genetic operators that work on solution coding to produce new individuals. One of the key issues of EA is the loss of diversity due to the selective pressure that leads to a *premature convergence* toward local optima. Since the early work on EA, the ability to maintain a good trade-off between exploration of the search space and exploitation of potential solutions has been investigated to avoid premature convergence. A classical solution resides in the promotion of *speciation* in artificial populations and various methods have been proposed. The underlying idea is most of the time based on the concept of mating restriction. For example, a distance measure between individuals can be introduced (genotypic or phenotypic) and the reproduction is controlled according the similarity between individuals (Golberg & Richardson 1987). An alternative is to impose a geographical context to the populations (islands or grids models) and then to limit the reproduction to spatially close individuals (Spiessens & Manderick 1991, Gorges-Schleuter 1991).

The cellular genetic algorithms (cGAs) are a subclass of EA in which the population is embedded upon a grid. Recently Simoncini et al. (2006) have extended the cGAs with the anisotropic selection. In this algorithm, a parameter controls the spatial anisotropy of the grid allowing the promotion of the diversity and an accurate regulation of the selective pressure during the evolutionary process. Studies in ecology have demonstrated that this kind of assortative mating should be enough to produce speciation (Almeida & de Abreu 2005).

This paper deals with the possibility of creating and maintaining species with spatial anisotropy in the cGAs and it is structured as follows: the section 2 starts with an overview of the cGAs then it explains how the anisotropic selection is implemented and finally the influence of anisotropy in both selective pressure and algorithm performance is reported. In section 3, the classical definitions of speciation in nature are presented and

the evolutionary dynamics of cGA with anisotropy selection is compared to parapatric speciation.

## 2 Structured populations and Spatial anisotropy

### 2.1 Cellular Genetic Algorithms

The cellular Genetic Algorithms (cGAs) are a subclass of Evolutionary Algorithms in which the population is generally embedded on a two dimensional toroidal grid. In this kind of algorithms, exploration and population diversity are enhanced thanks to the existence of small overlapped neighborhoods (Spiessens & Manderick 1991). An individual of the population is placed on each cell of the grid and represents a solution of the problem to solve. An evolutionary process runs simultaneously on each cell of the grid, selecting parents from the neighborhood of the cells and applying operators for recombination, mutations and replacement for further generations. Such a kind of algorithms is especially well suited for complex problems (Jong & Sarma 1995). One of the interests of cGA is to slow down the convergence of the population among a single individual. Complex problems often have many local optima, so if the best individual spreads too fast in the population it will improve the chances to reach a local optimum of the search space. Slowing down the convergence speed can be done by slowing down the selective pressure on the population.

### 2.2 Anisotropic selection

The anisotropic selection is a selection method in which the neighbors of a cell may have different probabilities to be selected. The Von Neumann neighborhood of a cell  $C$  is defined as the sphere of radius 1 centered at  $C$  in Manhattan distance. The anisotropic selection assigns different probabilities to be selected to the cells of the Von Neumann neighborhood according to their position and so corresponds to a spatial anisotropy. The probability to choose the center cell  $C$  remains fixed at  $\frac{1}{5}$ . Let us call  $p_{ns}$  the probability of choosing the cells North ( $N$ ) or South ( $S$ ) and  $p_{ew}$  the probability of choosing the cells East ( $E$ ) or West ( $W$ ). Let  $\alpha \in [-1; 1]$  be the control parameter that will determine the probabilities  $p_{ns}$  and  $p_{ew}$ . This parameter will be called the *anisotropic degree*. The probabilities  $p_{ns}$  and  $p_{ew}$  can be described as:

$$p_{ns} = \frac{(1 - p_c)}{2}(1 + \alpha)$$

$$p_{ew} = \frac{(1 - p_c)}{2}(1 - \alpha)$$

Thus, when  $\alpha = -1$  we have  $p_{ew} = 1 - p_c$  and  $p_{ns} = 0$ . When  $\alpha = 0$ , we have  $p_{ns} = p_{ew}$  and when  $\alpha = 1$ , we have  $p_{ns} = 1 - p_c$  and  $p_{ew} = 0$ . Figure 1 shows a Von Neumann Neighborhood with the probabilities to select

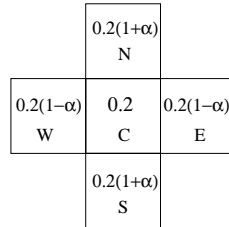


Figure 1: Von Neumann neighborhood with probabilities to choose each neighbor

each cell as a function of  $\alpha$ .

The anisotropic selection operator works as follows. For each cell it selects  $k$  individuals in its neighborhood ( $k \in [1; 5]$ ). The  $k$  individuals participate to a tournament and the winner replaces the old individual if it has a better fitness or with probability 0.5 if the fitnesses are equal. When  $\alpha = 0$ , the anisotropic selection is equivalent to a standard tournament selection and when  $\alpha = 1$  or  $\alpha = -1$  the anisotropy is maximal and we have an uni-dimensional neighborhood with three neighbors only. In the following, considering the grid symmetry we will consider  $\alpha \in [0; 1]$  only: when  $\alpha$  is in the range  $[-1; 0]$  making a rotation of  $90^\circ$  of the grid is equivalent to considering  $\alpha$  in the range  $[0; 1]$ .

Accounting for the value of  $\alpha$ , the anisotropic selection may induce an assortative mating. When the anisotropic degree is null, there is no anisotropy in selection method, the algorithm corresponds to the standard cellular GA. When the anisotropic degree is maximal, selection is computed between individuals in the same column only, the algorithm is then an island model where each sub-population is a column of the grid structured as a ring of cells with no interactions between sub-populations. When the anisotropic degree is set between low

and maximum value, according to selection, a number of individuals can be copied from one sub-population (i.e. column) to the adjacent columns.

### 2.3 Selective pressure

The selective pressure is related to the population diversity in cellular genetic algorithms. One would like to slow down the selective pressure when trying to solve multimodal problems in order to prevent the algorithm from converging too fast upon a local optimum. On the opposite side, when there is no danger of converging upon a local optimum, one would like to increase the selective pressure in order to obtain a good solution as fast as possible. A common analytical approach to measure the selective pressure is the computation of the takeover time (Rudolph. 2000) (Sprave. 1999). It is the number of generations needed for the best individual to conquer the whole grid when the only active operator is the selection (Goldberg & Deb 1990). Figure 2 shows the influence of the anisotropic degree on the takeover time. This figure represents the average takeover times observed on 1000 runs on a  $32 \times 32$  grid for different anisotropic degrees. It shows that the selective pressure is decreasing while increasing anisotropy. These results confirm that the anisotropic selection gives to the algorithm the ability to control accurately the selective pressure. They are fairly consistent with our expectation that selection intensity decreases when the anisotropic degree increases. However, the correlation between takeover and anisotropy is not linear; it fast increases after the value  $\alpha = 0.9$ .

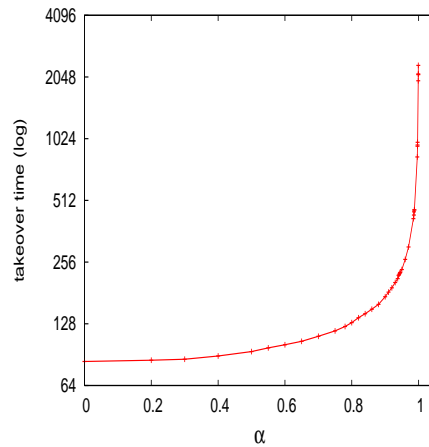


Figure 2: Average of the takeover time as a function of the anisotropic degree  $\alpha$ .

### 2.4 A case study : The quadratic assignment problem

We experimented the anisotropic selection on a Quadratic Assignment Problem (QAP): Nug30. The QAP is an important problem in theory and practice as well. It was introduced by Koopmans and Beckmann in 1957 and is a model for many practical problems (Koopmans & Beckmann 1957). The QAP can be described as the problem of assigning a set of facilities to a set of locations with given distances between the locations and given flows between the facilities. The goal is to place the facilities on locations in such a way that the sum of the products between flows and distances is minimal. Given  $n$  facilities and  $n$  locations, two  $n \times n$  matrices  $D = [d_{ij}]$  and  $F = [f_{kl}]$  where  $d_{ij}$  is the distance between locations  $i$  and  $j$  and  $f_{kl}$  the flow between facilities  $k$  and  $l$ , the objective function is:

$$\Phi = \sum_i \sum_j d_{p(i)p(j)} f_{ij}$$

where  $p(i)$  gives the location of facility  $i$  in the current permutation  $p$ . Nugent, Vollman and Ruml proposed a set of problem instances of different sizes noted for their difficulty (C.E. Nugent 1968). The instances they proposed are known to have multiple local optima, so they are difficult for a genetic algorithm. We experiment our algorithm on the 30 variables instance called Nug30.

#### 2.4.1 Setup

We use a population of 400 individuals placed on a square grid ( $20 \times 20$ ). Each individual represents a permutation of  $\{1, 2, \dots, 30\}$ . We need a special crossover that preserves the permutations:

- Select two individuals  $p_1$  and  $p_2$  as genitors.
- Choose a random position  $i$ .
- Find  $j$  and  $k$  so that  $p_1(i) = p_2(j)$  and  $p_2(i) = p_1(k)$ .
- exchange positions  $i$  and  $j$  from  $p_1$  and positions  $i$  and  $k$  from  $p_2$ .
- repeat  $n/3$  times this procedure where  $n$  is the length of an individual.

This crossover is an extended version of the UPMX crossover proposed in (Migkikh, Topchy, Kureichik & Tetelbaum n.d.). The mutation operator consist in randomly selecting two positions from the individual and exchanging those positions. The crossover rate is 1 and we do a mutation per individual. We perform 500 runs for each anisotropic degree. Each run stops after 1500 generations.

## 2.5 Best results

Figure 3 shows the average performance of the algorithm towards  $\alpha$  on the QAP: for each value of  $\alpha$  we average the best solution of each run. The purpose here is to minimize the fitness function values. The performances

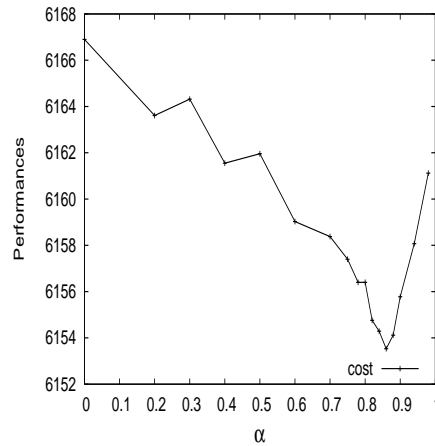


Figure 3: Average costs as a function of  $\alpha$  for the QAP.

are growing with  $\alpha$  and then fall down as  $\alpha$  is getting closer to its limit value. The best average performance is achieved for  $\alpha = 0.86$ . This threshold probably corresponds to a good exploration/exploitation trade-off: the algorithm favors propagation of good solutions in the vertical direction with few interactions on the left or the right sides. This kind of dynamics is well adapted to this multi-modal problem as we can reach local optima on each columns of the grid and then migrate them horizontally to find new solutions. The worst average performance is observed for  $\alpha = 0$ .  $\alpha = 0.86$  corresponds to the optimal trade-off with the best migration rate between sub-populations. In our model, the migration rate is not the number of individuals which are swap between sub-populations, but the probability for the selection operator to choose two individuals from separate columns: two individuals from separate sub-populations would then share information. We can tell that there is an optimal migration rate that is induced by the value of the anisotropic degree  $\alpha$ . Performances would probably improve if the migration rate did not stay static during the search process. As in (Alba & Dorronsoro 2005), we can define some criteria to self-adjust the anisotropic degree along generations.

## 3 Speciation and Spatial anisotropy

### 3.1 Definitions

The two most discussed types of speciation found in the literature are the sympatric and the allopatric speciations. The allopatric speciation is due to genetic drift and natural selection in two populations separated by a geographic barrier. Along several generations, the two populations develop so many differences that they become reproductively isolated and can no longer interbreed. The sympatric speciation is far more complex since in this case there is no physical barrier to prevent gene flow. Non random mating and competition are the essential ingredients (Gavrilets & Gravner. 2000), although some authors claim that assortative mating alone is enough to produce reproductive isolation followed by sympatric speciation (Almeida & de Abreu 2005).

The parapatric speciation can be seen as the most general case for speciation that includes both allopatric and sympatric speciation (Gavrilets 2004). The population occupies a spatially continuous habitat and adaptation evolves from a gradient which may or not result in speciation. We think this latter scenario can be reproduced in anisotropic cGAs .

### 3.2 Emergence of Species

In this section, the polymorphism of the population is used to monitor the emergence of species. We study the evolution of the genotypic diversity during runs within either the whole grid, the rows and the columns. For that purpose, three measures on the population diversity are observed along generations on 100 independent runs for each anisotropic degree.

First, we define the global population diversity  $gD$ :

$$gD = \left(\frac{1}{\#r\#c}\right)^2 \sum_{r_1, r_2} \sum_{c_1, c_2} d(x_{r_1 c_1}, x_{r_2 c_2})$$

where  $d(x_1, x_2)$  is the distance between individuals  $x_1$  and  $x_2$ . The distance is inspired from the Hamming distance: It is the number of locations that differs between two individuals divided by their length  $n$ . Then, we measure the diversity inside potential sub-populations (vertical diversity) and diversity between sub-populations (horizontal diversity). The vertical (resp. horizontal) diversity is the sum of the average distance between all individuals in the same column (resp. row) divided by the number of columns (resp. rows):

$$vD = \frac{1}{\#r} \frac{1}{\#c^2} \sum_r \sum_{c_1, c_2} d(x_{r c_1}, x_{r c_2})$$

$$hD = \frac{1}{\#c} \frac{1}{\#r^2} \sum_c \sum_{r_1, r_2} d(x_{r_1 c}, x_{r_2 c})$$

where  $\#r$  and  $\#c$  are the number of rows and columns in the grid.

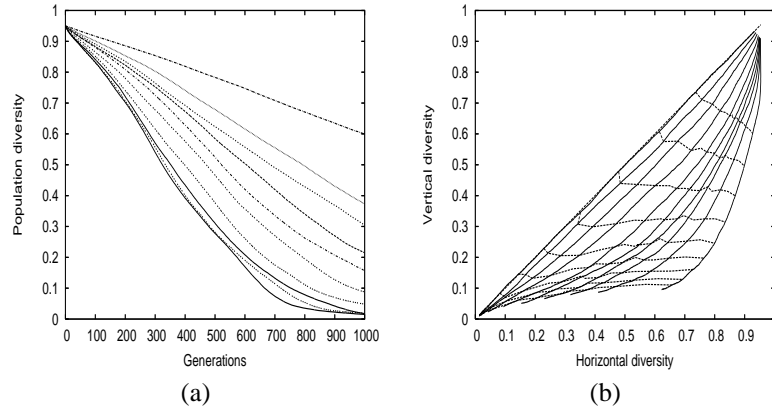


Figure 4: Global population diversity against generation, with increasing  $\alpha$  from bottom to top (a) and vertical diversity against horizontal diversity, with increasing  $\alpha$  from left to right (b).

Figure 4(a) shows the average global diversity observed on the 1000 first generations during 100 runs on the QAP. The curves from bottom to top correspond to increasing values of  $\alpha$  from zero to nearly one. Experiments measuring genetic diversity show that small migration rate ( $\alpha$  close to one) causes sub-population to dominate others and retain global diversity without being able to exchange solutions to produce better results. This case corresponds to a form of allopatric speciation define where a barrier prevents genetic material to be exchanged between close individuals. At the opposite, with  $\alpha$  closed zero, the global diversity falls near to zero after 800 generations causing premature convergence and negatively affects performances (see figure 3).

Figure 4(b) represents the vertical diversity against the horizontal diversity. The contour lines plotted every 100 generations give some information on the speed of decrease of diversity. The more the migration rate decreases (i.e.  $\alpha$  increases), the more the diversity is maintained on each row and sub-populations converge in each column. The vertical and horizontal diversities are decreasing with the same speed for the cellular model ( $\alpha = 0$ ) and

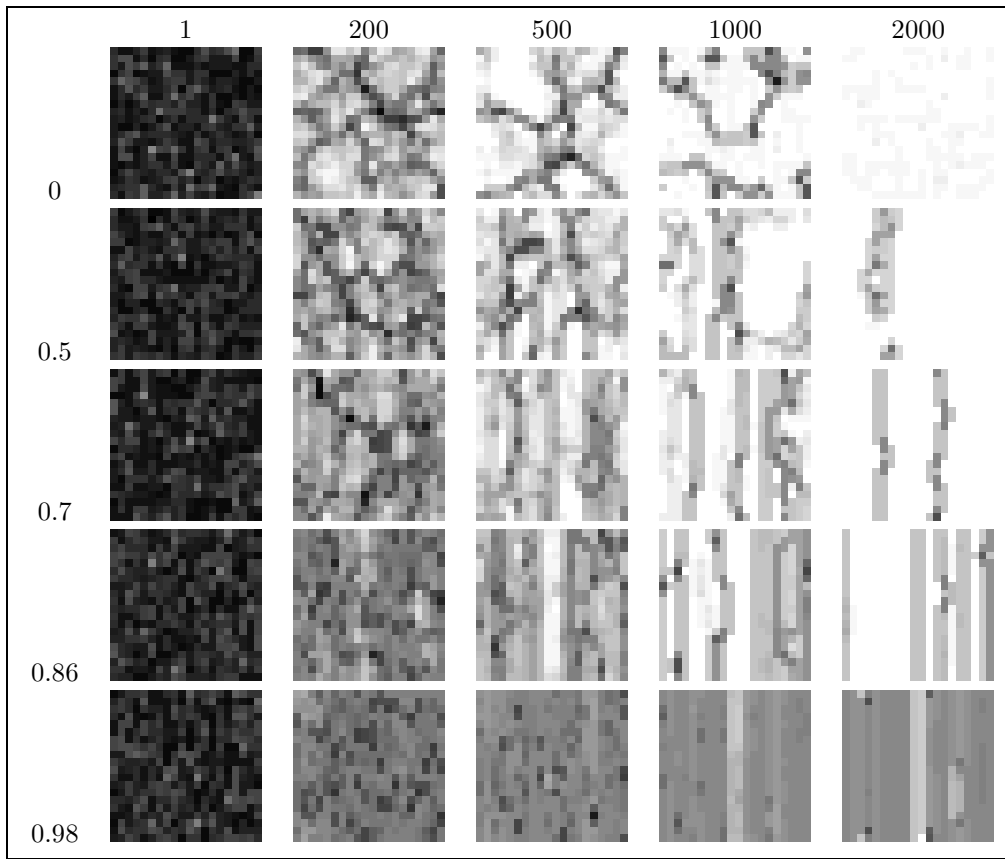


Figure 5: Local diversity in the population along generations (left to right) for increasing  $\alpha$  (top to bottom)

lower number of interactions between sub-populations helps the algorithm to maintain diversity on the rows when  $\alpha$  is high.

Figure 5 shows snapshots of the population diversity during one single run at different generations. The snapshots are taken from left to right at generations 1, 200, 500, 1000 and 2000. The parameter  $\alpha$  takes values in  $\{0, 0.5, 0.7, 0.86, 0.98\}$  from top to bottom. Each snapshot shows the genotypic diversity in the neighborhoods of all cells on the grid. Color black means maximum diversity and color white means that there is no more diversity in the cell's neighborhood. Those snapshots help to understand the influence of the anisotropic selection on the genotypic diversity. First, we can see that the anisotropic degree influences the dynamic of propagation of good individuals on the grid. This propagation is the cause of the loss of diversity in the population. In the standard cellular model ( $\alpha = 0$ ), good individuals propagate roughly circularly, in both directions. If we slightly privilege the vertical direction ( $\alpha = 0.5$ ) the circles become elliptical. As  $\alpha$  increases, the dynamic changes and good individuals propagate column by column. For extreme values of the anisotropic degree ( $\alpha$  close to 1) the migration rate is so low that good individuals are stuck in the sub-populations and the sharing of genetic information with other sub-populations is seldom observed. In that case, the selective pressure is too low and it negatively affects performances. The crossover operator doesn't have any effect in the white zones, since they represent cells with no more diversity in their neighborhoods. For the standard cellular case, interactions between cells may have some effects on performances only at the frontier between the circles. It represents a little proportion of cells on the grid after a thousand generations. For  $\alpha = 0.86$ , we can see vertical lines of diversity, which means that good individuals appear in each sub-populations. For example, when we see two adjacent columns colored in gray, it means that in those columns, two different species have emerged. At generation 2000, a good individual has colonized the left of the grid but he still can share information with individuals in the gray zones. This is clearly an example of speciation in the presence of a (small) gene flux. We think that with such a setting the propagation of the genetic information through the whole grid is possible.

### 3.3 Maintenance of Species

To show up to what extent anisotropic selection allows to maintain speciation, we have conducted experiments where two different individuals with the best fitness (here 1) are placed on a  $64 \times 64$  square grid at the initial

generation. These optimal individuals are spatially distant from each other in the less favored direction (here oriented horizontally). Figure 6 shows some snapshots of the spreading of these two best over generations for different anisotropic degrees. Cells in light gray (resp. dark gray) are copies of the first best (resp. the second best), and all white cells have a null fitness value. Generations grow from top to bottom, and the anisotropic parameter  $\alpha$  increases from left to right. The left-hand row ( $\alpha = 0$ ) represents standard binary tournament schema ; we observe that standard selection is not able to maintain species, after 1000 generations the grid is a mixture of the two optima. On the other hand, as  $\alpha$  increases, two stable frontiers between species emerge. Here, the spatial anisotropy increases *cohesion* in each cells lineages.

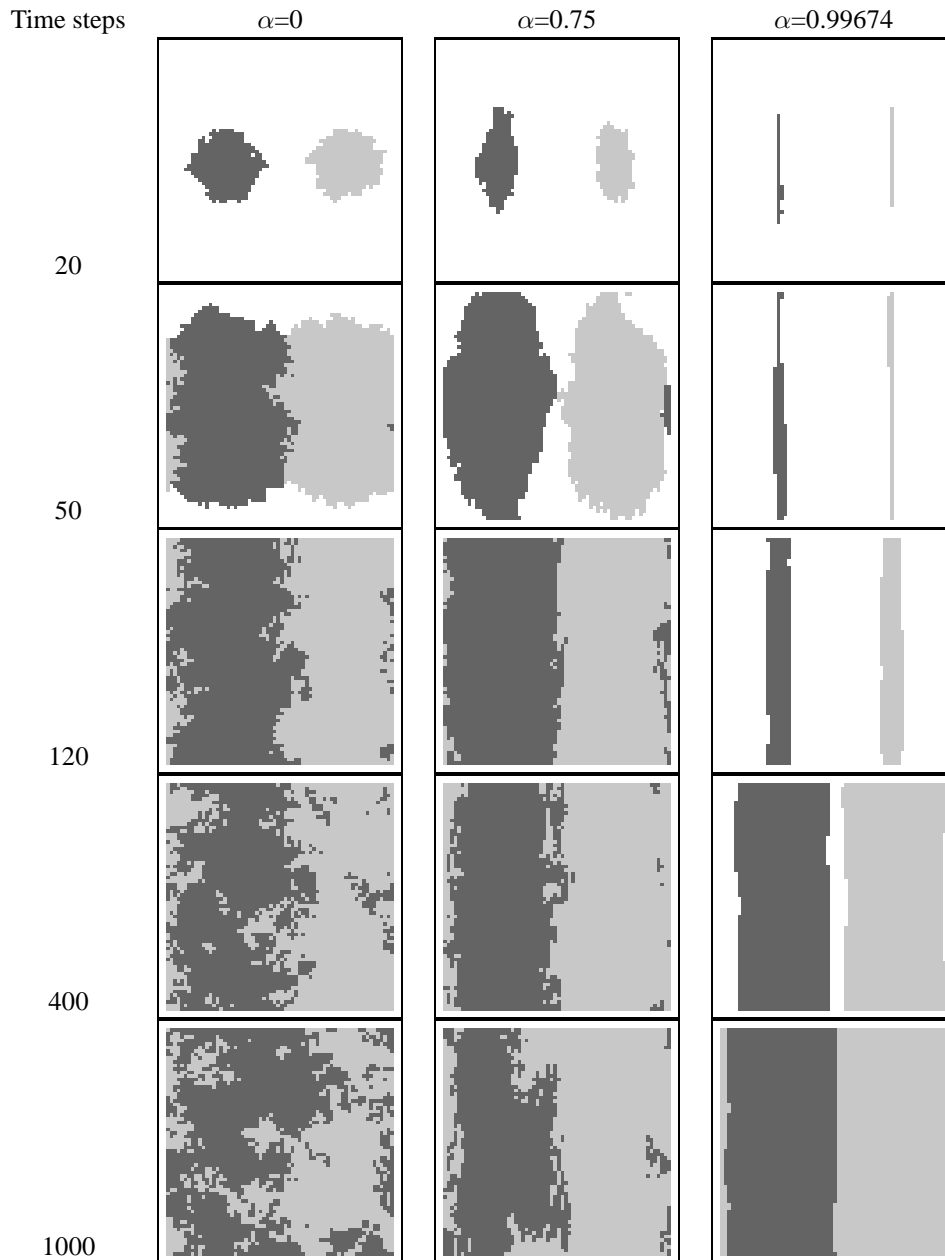


Figure 6: Spreading of two species

#### 4 Conclusion and Perspectives

In cellular Genetic Algorithms (cGAs), the populations are structured within a grid. In this paper, the spatial anisotropy of the neighborhood of each cell can promote the mating between individuals according to a given specific direction. On the quadratic assignment problem, the anisotropic degree plays a major role with regard to the average fitness found. Indeed, the performances increases with the anisotropic degree  $\alpha$  until a threshold

value ( $\alpha = 0.86$ ). After this threshold, an artificial barrier is created between sub-populations, and the migration rate between columns becomes too small to generate good solutions.

An empirical study of the local diversity in cGAs illustrates the variations of the interactions between cells for different settings of the system. The propagation of genetic material, which is strongly related to the genotypic diversity in the population, depends on the anisotropic degree. Propagation of good individuals looks like circles for low values of  $\alpha$  and turns to vertical lines for high values of  $\alpha$ . The genotypic diversity is maintained in the population when the anisotropic degree is high, but when it reaches extreme values, the few sporadic interactions between columns penalize the performances of the algorithm. In the latter case, the emergence of species is intuitively expected. However, in theory, the presence of a small gene flux between different populations does not prevent them from speciation. In this paper, we have reproduced such behaviors with the anisotropic selection in cGAs, since the possibility of evolving and maintaining species, while preserving a continuous exchange of information is empirically demonstrated.

## References

- Alba, E. & Dorronsoro, B. (2005). "The Exploration/Exploitation Tradeoff in Dynamic Cellular Genetic Algorithms" *IEEE transactions on Evolutionary Computation*. Vol. 9(2). pp. 126–142.
- Almeida, C. R. & de Abreu, F. V. (2005). "Dynamical instabilities lead to sympatric speciation" *Evolutionary Ecology Research*. Vol. 5. p. 730.
- C.E. Nugent, T.E. Vollman, J. R. (1968). "An Experimental Comparison of Techniques for the Assignment of Techniques to Locations" *Operations Research*. **16**: 150–173.
- Gavrilets, S. (2004). *Fitness landscapes and the origin of species*. Princeton University Press.
- Gavrilets, S, R. A. & Gravner, J. (2000). "Dynamics of speciation and diversification in a metapopulation" *Evolution*. **54**: 1493–1501.
- Golberg, D. E. & Richardson, J. (1987). "Genetic algorithms with sharing for multimodal function optimization" *International Conference on Genetic Algorithms*. pp. 41–49.
- Goldberg, D. E. & Deb, K. (1990). "A Comparative Analysis of Selection Schemes Used in Genetic Algorithms" *FOGA*.. pp. 69–93.
- Gorges-Schleuter, M. (1991). "Explicit Parallelism of Genetic Algorithms through Population Structures" *PPSN*.. pp. 150–159.
- Jong, K. A. D. & Sarma, J. (1995). "On Decentralizing Selection Algorithms." *ICGA*.. pp. 17–23.
- Koopmans, T. & Beckmann, M. (1957). "Assignment Problems and the Location of Economic Activities" *Econometrica*. **25**(1): 53–76.
- Migkikh, V. V., Topchy, A. P., Kureichik, V. M. & Tetelbaum, A. Y. (n.d.). "Combined Genetic And Local Search Algorithm For The Quadratic Assignment Problem".  
**URL:** [citeseer.ist.psu.edu/78722.html](http://citeseer.ist.psu.edu/78722.html)
- Rudolph, G. (2000). "On takeover times in spatially structured populations: Array and ring." *Proceedings of the Second Asia-Pacific Conference on Genetic Algorithms and Applications (APGA '00)*.. pp. 144–151.
- Spiessens, P. & Manderick, B. (1991). "A Massively Parallel Genetic Algorithm: Implementation and First Analysis" *ICGA*.. pp. 279–287.
- Sprave, J. (1999). "A unified model of non-panmictic population structures in evolutionary algorithms." *Proceedings of the congress on Evolutionary computation*.. pp. 1384–1391.