

Spatial Embedding and Loss of Gradient in Cooperative Coevolutionary Algorithms

R. Paul Wiegand* and Jayshree Sarma

Department of Computer Science, George Mason University, Fairfax, VA 22030, USA
paul@tesseract.org, jsarma@cs.gmu.edu

Abstract. Coevolutionary algorithms offer great promise as adaptive problem solvers but suffer from several known pathologies. Historically, spatially embedded coevolutionary algorithms seem to have succeeded where other coevolutionary approaches fail; however, explanations for this have been largely unexplored. We examine this idea more closely by looking at spatial models in the context of a particular coevolutionary pathology: loss of gradient. We believe that loss of gradient in cooperative coevolution is caused by asymmetries in the problem or initial conditions between populations, driving one population to convergence before another. Spatial models seem to lock populations together in terms of evolutionary change, helping establish a type of dynamic balance to thwart loss of gradient. We construct a tunably asymmetric function optimization problem domain and conduct an empirical study to justify this assertion. We find that spatial restrictions for collaboration *and* selection can help keep population changes balanced when presented with severe asymmetries in the problem.

1 Introduction

Coevolutionary algorithms (CEAs) are increasingly popular extensions of traditional evolutionary algorithms (EAs). The most fundamental differences between CEAs and EAs stem from the adaptive nature of fitness evaluation in coevolutionary systems: individuals are assigned fitness values based on direct interactions with other individuals. Examples of such systems include competitive approaches where an individual in one population receives fitness based on the result of a competition with one or more individuals, as well as cooperative approaches where an individual represents a component of a larger, more structured problem, and receives fitness based on how well it performs in conjunction with individuals from other populations.

There is a lot of intuitive appeal to coevolutionary algorithms. In the case of competitive systems, there is the hope of establishing an *arms race*, where steady progress is made by mutual and reciprocal adaptations between competing groups of individuals [1]. Cooperative systems have the same sort of hope, though perhaps the term “parallel adaptive changes” might be more appropriate to its collaborative nature. In spite of their appeal, coevolutionary algorithms are often challenged by seemingly simple problems. Constructing algorithms that facilitate arms-race type behaviors is far from easy.

* R. Paul Wiegand currently holds a postdoctoral position with the American Society for Engineering Education and conducts research at the Naval Research Laboratory.

One difficulty common to both competitive and cooperative coevolutionary algorithms is the *loss of gradient* problem, in which one population comes to severely dominate the others, creating an impossible situation where the other populations have insufficient information from which to learn. Suppose a grand master of chess repeatedly plays a child, who is merely learning: if the child gets no more information than the result of the game, he is unlikely to learn much from the process.

Of particular interest to us are cooperative coevolutionary algorithms (CCEAs). These algorithms have appeal when applied to problems with large domain spaces having certain structural properties among interacting components. The intuition behind this advantage is that the algorithm adaptively searches only projections of the space at any given time, thus presenting a narrower search domain in a particular generation. However, loss of gradient can also occur in CCEAs when the diversity of a subset of the populations suddenly decreases, leaving the others searching only a static projection and not the full problem [2].

The key to solving the loss of gradient problem would seem to be found in helping the algorithm maintain some kind of balance between the populations in terms of evolutionary change. Here biology suggests a potential solution: establish some locality constraints in terms of selection and collaborator interactions [3]. Indeed there is precedent for such fine-grained spatially embedded population models in both traditional EAs [4], as well as coevolutionary systems [5]. Moreover, recent empirical research has suggested that spatial models may indeed give CEAs some kind of advantage, though the reason for this advantage has remained unexplored. We believe that such spatial CEA successes are due to their ability to thwart the loss of gradient problem by constraining the speeds at which populations themselves can propagate changes, as well as the speeds at which populations can provide interaction information to the other populations. This paper explores the relationship between spatial models and the causes of loss of gradient in cooperative coevolutionary algorithms. We make no claims regarding the general utility of spatial CCEAs for solving static function optimization problems, rather we seek only to expose some of the reasons why spatial embedding may benefit coevolution when solving problems that have certain properties.

The next section provides background of the cooperative coevolutionary framework we consider, the loss of gradient pathology, and spatial embedding in EAs. The third section describes our spatial CCEA, including the effects of constraining the locality of certain operators. Our experimental design and results are described in detail in the fourth section. The paper terminates with a section discussing our conclusions, as well as indicating areas of future work highlighted by our research.

2 Background

There are a variety of models of cooperative coevolution, beginning with the early work of Husbands and Mill [6], to more recent models by Potter [7], as well as Moriarty and Miikkulainen [8]. This paper focusses on the Potter model of cooperative coevolution. Here each population contains individuals that represent a particular component of the problem, so that one member from each population is needed in order to assemble a complete solution. Evaluation of an individual from a particular population is performed by assembling the individual with collaborating partners from other pop-

ulations. Often multiple collaborators are used to gain a better quality estimate of an individual's contribution to the interaction. An individual's fitness could be the mean of such evaluations, or the max, among other approaches. Aside from evaluation, the populations are evolved independently.

An example may help clarify things. Suppose we are optimizing a two argument function, $f(x, y)$. One might assign individuals in the first population to represent the x argument and the second to represent y . Each population is evolved separately, except that when evaluating an individual in some population (e.g., x), collaborating representatives must be chosen from the other population (y) in order to obtain an objective function value with a complete solution, $f(x, y)$. A simple example collaboration method is to choose a representing member by using the most fit individual from the other population as determined by the previous round of evaluations. Another approach is to pick partners at random from the other population. Once a complete solution is formed, it can be evaluated and the resulting score can be assigned to the individual as the mean of several of these interactions.

Though the behavior of cooperative and competitive CEAs can certainly differ, they also share some pathologies. One such pathology is that of loss of gradient. In a multi-population model, loss of gradient occurs when one population converges to strategies that provide the other populations no hope for meaningful search due to the unavailability of informational distinctions between individuals. In a competitive system, this suggests that one population has severely dominated the other, such that no information is learned from a contest between individuals from those populations. In a cooperative system, this suggests that one or more populations has converged, and the projection offered by these converged populations during collaboration is misleading or degenerate in some way for the populations that are still attempting to progress.

The term *loss of gradient* stems primarily from three works relating to the analysis of competitive coevolutionary algorithms. Juillé and Pollack [9], as well as Ficici and Pollack [10], primarily focus on methods for measuring and maintaining coevolutionary progress while discussing the need for successful competitive algorithms to maintain a gradient of search. Watson and Pollack [11] specifically introduce the term, using a simple problem structure to help illustrate and identify it as one of several pathologies of competitive coevolutionary systems. While more recent work suggests that loss of gradient may not be as big a problem for competitive algorithms as problems related to overspecialized focussing [12], it seems evident that it remains a significant challenge. Multi-population coevolution works by making parallel adaptive changes in interacting populations, but a balance between these changing populations must exist in order maintain co-adaptive search gradients. When that balance is lost, the search can fail by forcing changes in the populations to become disengaged, resulting in polarization of the populations in terms of subjective fitness assessment [11].

Though loss of gradient is perhaps more easily understood in the context of competitive coevolution, the same problem challenges cooperative models. It can happen, for instance, when rapid asymmetric changes in one population lead it to converge to near homogeneity, forcing the other disengaged populations to be driven into arbitrary equilibria [2]. Such inequality can be created by many factors including initialization effects, asymmetries in algorithmic choices or asymmetries in the problem itself.

Biologists and social scientists have studied spatially embedded evolution and co-evolution for some time, typically taking game-theoretic approaches to understanding the dynamics in single population systems playing games such as iterated prisoner’s dilemma [13]. Researchers in the field of evolutionary computation have also studied spatial models in some detail [4, 14], but analysis of coevolutionary spatial models has primarily been relegated to discussions of general CEA performance measures [1]. Nevertheless, applications of such systems have proved effective [15–17], sometimes demonstrating clear advantages over non-spatial CEAs [5]. Unfortunately, the underlying cause of this advantage has not been addressed.

3 A Spatial CCEA

The basic idea behind spatially embedded evolutionary algorithms is quite simple. In the most obvious case, the individuals in the population are distributed on a 2-D (often toroidal) grid. Each individual occupies a specific location on that grid, and locality is defined by the vertical and horizontal topological connections. The algorithm works the same as a traditional evolutionary algorithm, with the exception of selection. New individuals for each position are selected locally using a pre-defined neighborhood. Virtually any traditional selection operator can be applied in this method, though its purview is limited to the individuals in the local neighborhood of the position under consideration. Updates to the current population may be synchronous or asynchronous.

For example, consider a synchronous spatial variant of a traditional generational evolutionary algorithm. Let’s suppose the algorithm uses binary representation with bit-flip mutation, fitness proportionate selection, and a 2×2 , von Neumann neighborhood (a diamond shaped subset of the grid covering five points). Individuals are distributed on the grid positions, initialized uniformly at random, then evaluated against the objective function. We then produce a new child for each grid position by selecting from the individuals in the surrounding neighborhood, proportionally according to their fitness values. The offspring produced for that position is subjected to mutation and placed into the next generation’s grid at the same position. A new position is then considered in the same way, and the process is repeated until all positions have been considered. At this point the generation is complete and a new generation begins.

The most distinctive parameter of spatial models of this sort is the neighborhood definition. It turns out that for several geometric shapes, a general size measure (called the *radius* of the neighborhood) is useful for understanding diffusive properties of the selection method. The larger the radius, the faster information is propagated throughout the population, the smaller the radius, the slower [4].

In the case of a simple, spatial CCEA, things are very similar. We now consider separate grids for each population, and “align” the 2-D grids in a stacked fashion, forming a 3-D lattice. Now a given population grid position has an adjacent position above or below the current position in another population grid. In addition to selection, collaboration now makes use of the neighborhood mechanism. Representative partners are selected from a neighborhood in the adjacent population(s), providing another type of locality restriction altogether—one that controls the amount of information about interactions a population receives. In this case the larger the radius of the collaboration

neighborhood, the more information about potential partners is available, the smaller the radius, the less information.

Controlling these two relative sizes appears to be very important to the success of coevolution in many cases. By restricting the rate of diffusion due to selection, the rates of change of EAs associated with each population can be reduced. By restricting the amount of available interaction information, the EAs can be given time to adapt to radical phenotypic shifts in the collaborating population before being exposed to degenerate partners. If rapid, asymmetric losses of diversity are our major concern then local selection slows the rate at which populations reach homogeneity and local collaboration allows topologically distant partners the chance to make changes before being exposed to the new individuals. The intuition is that the overall effect acts as a kind of cap to population advancement, “locking” the populations together in terms of their relative evolutionary changes.

4 Experimental Results

4.1 The ASYMMTWOQUAD Problem

Recall that asymmetries in a problem can exacerbate the loss of gradient pathology for the CCEA. To help show this, we define a class of problems that allows one to construct instances that vary in terms of degree of problem asymmetry. The problem class defined below consists of two quadratic peaks, one a global maxima and the other a local submaxima. Each of these two peaks can be varied independently to adjust asymmetric conditions in the landscape between the two arguments.

Definition 1. Given constant values k_1 and k_2 defining the relative heights of two peaks; parameters s_{x1} , s_{x2} , s_{y1} , and s_{y2} defining peak widths; and points (\bar{x}_1, \bar{y}_1) and (\bar{x}_2, \bar{y}_2) defining the locations of the peaks, the function $\text{ASYMMTWOQUAD} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ (ATQ) is defined by

$$\begin{aligned} \text{quad}_1(x, y) &= k_1 - \left[s_{x1} \cdot (\bar{x}_1 - x)^2 + s_{y1} \cdot (\bar{y}_1 - y)^2 \right] \\ \text{quad}_2(x, y) &= k_2 - \left[s_{x2} \cdot (\bar{x}_2 - x)^2 + s_{y2} \cdot (\bar{y}_2 - y)^2 \right] \end{aligned}$$

$$\text{ASYMMTWOQUAD}(x, y) = \max(\text{quad}_1, \text{quad}_2)$$

For our purposes, the parameters controlling the width of the peaks are most salient. The larger the s_{x1} value is, the narrower the first peak becomes along the x -axis. We can create a situation that endangers the CCEA of suboptimal convergence by tightening the global peak’s width relative to the suboptimal peak [2], and we can make this situation increasingly asymmetric by tightening more along one axis than another. Here, the s_{x2} and s_{y2} parameters were both set to 1, and the s_{x1} parameter was held fixed at 8. The s_{y1} parameter, however, was varied using the values $\{8, 16, 32, 64, 128, 256, 512\}$. The two optima are located at $(8, 1)$ and $(1, 8)$, and their objective values (k_1 and k_2) are 180 and 140, respectively. Domain values for x and y were restricted to $[1, 8]$. This problem class is simple enough to intuit salient properties, while allowing a researcher to generate a range of problems from very simple to quite difficult. Though the domain itself is relatively limited in scope, it serves us well here since our goal is to understand the effects spatial representations have on loss of gradient, not to demonstrate any specific advantage of spatial CCEAs over EAs in general.

4.2 Basic Experimental Setup

In all cases, the basic underlying algorithms were generational EAs with fitness proportionate selection. There were two populations, each containing individuals representing one of the two arguments to the ASYMMTWOQUAD function. The genotype of each individual was encoded as a binary string, each argument ranging in $[1, 8]$. The string length and population sizes were fixed for these experiments. There were 100 individuals in each population, each of which were $l = 128$ bits long, thus $n = 256$ bits. Collaboration was handled as follows. Five representatives from the alternate population were selected at random from either the entire population, or from the locally restricted areas defined by the spatial approaches. Each representative was evaluated with the current individual, and the average score was assigned for fitness purposes.

Though other population sizes, string lengths, and collaboration sizes are not shown, some were tested and the results were consistent with those reported here. It is our intent that the population size is linear with respect to the string length and that the number of collaborators is bounded above by the square root of the population size (with a constant factor of $1/2$). Such a rule leads to reasonably sized populations for the ASYMMTWOQUAD problem class and provides useful information for future analysts. Bit-flip mutation is used at a rate of $1/l$. There is evidence that crossover may exacerbate the effects of loss of gradient in CCEAs [2], and since our intent is to try to understand the effects of spatial embedding on loss of gradient at the most basic level we choose not to complicate the discussion in this article by including crossover in our discussion. This notwithstanding, though not shown, we also ran our experiments with parameterized uniform crossover operator of varying rates and the results are consistent with the findings stated here. Where possible, we keep things intentionally as simple as possible here to allow for future analysis.

4.3 Balancing Evolutionary Change

As we've already mentioned, the ASYMMTWOQUAD problem described above becomes increasingly more asymmetric as the s_{x1} parameter is increased. The effect of this is to make the global peak narrower in one dimension. This creates a situation in which one population has a clear advantage if for no other reason than the fact that the ratio of domain coverage of the global peak is larger in one dimension than another. The resulting effect on simple CCEAs is a decrease in performance due to loss of gradient.

Our hypothesis is that the locality constraints of selection and collaboration in the spatially embedded model will help keep the system in balance in terms of evolutionary change, thus improving performance. We first consider two algorithms, a more traditional non-spatial CCEA described above and a spatially embedded analog of this algorithm. In the spatial model, individuals in each population are spread out on a 2-D toroidal 10×10 grid. The neighborhood sizes for selection and collaboration are performed in a 2×2 diamond of radius 1 from a given position (covering 5 grid points).

To test these two algorithms we ran 14 experimental groups, 7 with a traditional non-spatial CCEA, and 7 with a spatial CCEA. These 7 groups correspond the ASYMMTWOQUAD problem parameter values $s_{x1} \in \{8, 16, 32, 64, 128, 256, 512\}$. The results are shown in two graphs in Figure 1. In both graphs, the x -axis indicates results from the different experimental groups specified by the s_{x1} parameter. Each point of the top graph illustrates the mean best-ever fitness value of 50 trials and the vertical

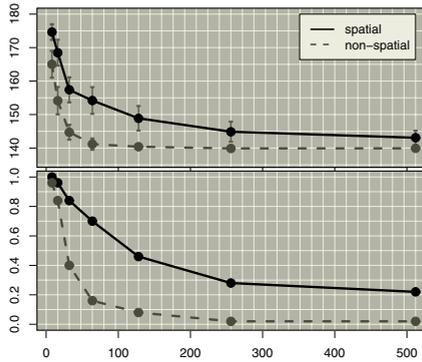


Fig. 1. Spatial and non-spatial CCEA results on the ASYMMTWOQUAD problem as s_{x1} (x -axis) varied. Points in the top graph represent mean and confidence intervals of best-ever (y -axis) values in 50 trials. Points in the bottom graph represent the ratio of trials (y -axis) that are likely to globally converge.

wings surrounding that point illustrate the 95% confidence intervals of that group. Each point in the bottom graph represents the ratio of the 50 trials where the best-ever value exceeded 140, and thus indicates the ratio of populations that plausibly may eventually converge to the global peak.

The best-ever values of the groups were tested using pair-wise t -tests with Bonferroni adjustment. The result indicates that for all values of s_{x1} the spatial model significantly outperforms the non-spatial model (with confidence of 95%). These results were consistent for population size of 49, $l = 64$ using 3 collaborators. They were also consistent with results applying parameterized uniform crossover with a crossover rate of 100% and swap rates of 0.5 and 0.2 (not shown).

The lower panel of the figure gives us some clue as to why the spatial groups resulted in higher performance. As the problem becomes increasingly asymmetric, the ratio of suboptimally converging populations drops much faster for the non-spatial CCEA than for the spatial CCEA.

Although it is not shown here for space reasons, it is also the case that the standard deviations in fitness values are significantly higher for the spatial model than for the non-spatial model in every case, indicating higher levels of diversity. Examining the inter-population diversity, the relative differences in standard deviations between populations of the same model, helps reveal something about the differing rates of change in the two populations. Here, the inter-population diversity within the first ten generations (the early stages of population convergence) reveals that the standard deviations of the two populations in the non-spatial CCEA differ significantly. This is only true for the $s_{x1} = 256$ case for the spatial model. This suggests that the spatial model may very well have an advantage with respect to differing rates of evolutionary change.

4.4 Collaboration and Selection

There are at least three ways this spatial model might improve upon the non-spatial model: local restrictions to selection, local restrictions to collaboration, and a combined impact of both of these.

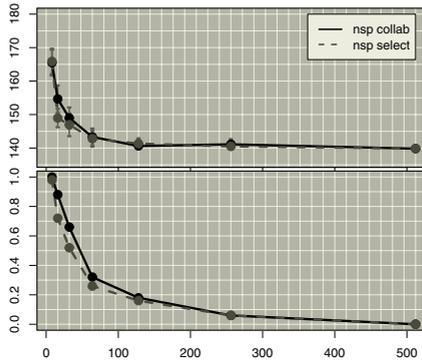


Fig. 2. Results for two CCEAs are presented: non-spatial selection with spatial collaboration and non-spatial collaboration with spatial selection. The results are reported for the ASYMMTWOQUAD problem as s_{x1} (x -axis) varied. Points in the top graph represent mean and confidence intervals of best-ever (y -axis) values in 50 trials. Points in the bottom graph represent the ratio of trials (y -axis) that are likely to globally converge.

First, an argument could be made that restricting selection to local neighborhoods cannot be responsible for the behavioral improvement. After all, the effects of locality on selection are mainly speed effects. One would expect to see a slowing of the pathology, but certainly not its absence. While this is true, it is also the case that spatial representations help maintain diversity, and this has been known to help resolve difficulties in certain kinds of acyclic problems [18].

The second difference between the spatial model and the non-spatial one is the locality restriction imposed upon collaboration. One may again be tempted to dismiss this component as the explanation for improvement since a smaller pool of potential collaborators seems to invite even more challenges to the CCEA. However, restricting collaborators to a local neighborhood restricts the flow of interaction information between the two populations, allowing one population to have more time with a given set of interactions before they are washed away by selection.

The most satisfying answer is that the improvement results from a combination of these two mechanisms. Slowing selection down and maintaining enough diversity to give one population time to catch up to the other, while restricting the information flow between the populations in such a way as to protect distant population members from unhelpful representatives from the collaborating population. We ran the same experiments as before for two more sets of groups, save that this time the spatial restrictions were used for only one of the two mechanisms. In the first set, selection is performed across the entire population but collaboration occurs in the 2×2 diamond neighborhood (“nsp select”), while in the second set the situation was entirely reversed (“nsp collab”). Figure 2 shows these results.

As can be seen from these graphs, the advantage the spatial model has over the non-spatial does indeed require *both* collaboration *and* selection to be restricted. Removing either produces results that are statistically indistinguishable from those of a completely non-spatial CCEA from the perspective of best-of-run performance results. Looking once again at the relative differences in standard deviations between populations in the

first ten generations is helpful. In almost all cases there were significant differences between population standard deviation scores, just as in the fully non-spatial case. This again suggests that both restrictions are necessary to explain the spatial CCEA's ability to thwart loss of gradient.

5 Conclusions and Future Work

Coevolution offers much promise, but achieving consistently successful performance from coevolutionary algorithms is a complicated and important research issue. Historically, many of the reported successes with coevolution have employed fine-grained, spatially embedded approaches. In some cases, researchers have reported success on a problem using a spatial embedding, while a distinct lack of success when a non-spatial CEA is applied to the very same problem [5]. This paper begins to answer why. The study of the general performance quality of spatial CCEAs is not germane here.

Recent analytical work has suggested that not only does the loss of gradient pathology affect cooperative coevolutionary algorithms, but it might be caused, or at least exacerbated by an imbalance in evolutionary change between co-adapting populations [2]. Our hypothesis here is that a spatial embedding is one way to help restore this balance by locking together the rates of relative changes between populations. We explore this hypothesis using a simple, tunably asymmetric landscape and use a cooperative coevolutionary algorithm to optimize this function.

We examined the CCEA with and without spatial restrictions for collaboration and selection together, as well as independently. The results showed that as the problem becomes increasingly asymmetric, the degradation in performance is far less in the fully spatial model than any other combination of spatially-restricted algorithms. Moreover, the ratio of trials that optimally converge was typically much higher for the spatial algorithm than for the non-spatial CCEA. Further, there seems to be some evidence that inter-population diversity measures are more disparate in the more non-spatial case. All of this leads us to conclude that spatially embedded cooperative coevolutionary algorithms use locality restrictions as a means of helping maintain balanced evolutionary change between populations, and that *both* the collaboration process *and* the selection process need to be so restricted. This effect appears to be largely due to the spatial model's ability to maintain larger levels of diversity in the population, but *also* because it keeps these levels somewhat balanced between populations.

The exact nature of these diversity differences has yet to be fully explored, and is a topic for further research. Additionally, a much more careful examination of the effects that locality restrictions have on competitive models, where loss of gradient is perhaps more intuitively understood, should be undertaken. Finally, as we develop a greater understanding for the causes of pathologies such as loss of gradient, we should begin exploring augmentations to traditional CEA approaches that help counteract these challenges. Spatial models are not the only solution to loss of gradient, but they are one more tool in a practitioner's toolbox to help them with such difficulties.

References

1. Cliff, D., Miller, G.F.: Tracking the red queen: Measurements of adaptive progress in co-evolutionary simulations. In: Proceedings of the Third European Conference on Artificial Life, Springer-Verlag (1995) 200–218
2. Wiegand, R.P.: An Analysis of Cooperative Coevolutionary Algorithms. PhD thesis, George Mason University, Fairfax, Virginia (2004)
3. Nuismer, S.L., Thompson, J.N., Gomulkiewicz, R.: Coevolution between hosts and parasites with partially overlapping geographic ranges. *Journal of Evolutionary Biology* **16** (2003) 1337–1345
4. Sarma, J.: An Analysis of Decentralized and Spatially Distributed Genetic Algorithms. PhD thesis, George Mason University, Fairfax, Virginia (1998)
5. Pagie, L.: Information Integration in Evolutionary Processes. PhD thesis, Universiteit Utrecht, Netherlands (1999)
6. Husbands, P., Mill, F.: Simulated coevolution as the mechanism for emergent planning and scheduling. In Belew, R., Booker, L., eds.: Proceedings of the Fourth International Conference on Genetic Algorithms, Morgan Kaufmann (1991) 264–270
7. Potter, M.: The Design and Analysis of a Computational Model of Cooperative CoEvolution. PhD thesis, George Mason University, Fairfax, Virginia (1997)
8. Moriarty, D., Mikkulainen, R.: Forming neural networks through efficient and adaptive coevolution. *Evolutionary Computation* **5** (1997) 373–399
9. Juillé, H., Pollack, J.: Coevolving the “ideal” trainer: Application to the discovery of cellular automata rules. In: Proceedings of the Third Annual Genetic Programming Conference, Madison, Wisconsin (1998)
10. Ficici, S., Pollack, J.: Challenges in coevolutionary learning: Arms-race dynamics, open-endedness, and mediocre stable states. In et al, A., ed.: Proceedings of the Sixth International Conference on Artificial Life, Cambridge, MA, MIT Press (1998) 238–247
11. Watson, R., Pollack, J.: Coevolutionary dynamics in a minimal substrate. In Spector, L., *et al*, eds.: Proceedings of the Genetic and Evolutionary Computation Conference (GECCO) 2001, Morgan Kaufmann (2001) 702–709
12. Bucci, A., Pollack, J.B.: Focusing versus intransitivity geometrical aspects of co-evolution. [19] 250–261
13. Nowak, M., May, R.: Evolutionary games and spatial chaos. *Nature* **359** (1992) 826–29
14. Giacobini, M., Alba, E., Tomassini, M.: Selection intensity in asynchronous cellular evolutionary algorithms. [19] 955–966
15. Hillis, D.: Co-evolving parasites improve simulated evolution as an optimization procedure. *Artificial Life II, SFI Studies in the Sciences of Complexity* **10** (1991) 313–324
16. Husbands, P.: Distributed coevolutionary genetic algorithms for multi-criteria and multi-constraint optimisation. In: Evolutionary Computing, AISB Workshop for Selected Papers, Springer-Verlag (1994) 150–165
17. Rongé, A., Nordahl, M.G.: Genetic programs and co-evolution developing robust general purpose controllers using local mating in two dimensional populations. In Voigt, H.M., Ebeling, W., Rechenberg, I., Schwefel, H.P., eds.: Parallel Problem Solving from Nature IV, Proceedings of the International Conference on Evolutionary Computation. Volume 1141 of LNCS., Berlin, Germany, Springer Verlag (1996) 81–90
18. Cartledge, J., Bullock, S.: Caring versus sharing: How to maintain engagement and diversity in coevolutionary populations. In: Proceedings from Seventh International Conference on Evolutionary Computation and Artificial Life. (2003)
19. Cantú-Paz, E., *et al*, eds.: Proceedings of the Genetic and Evolutionary Computation Conference (GECCO) 2003. In Cantú-Paz, E., *et al*, eds.: Proceedings of the Genetic and Evolutionary Computation Conference (GECCO) 2003, Berlin, Germany, Springer (2003)