

Sexual Recombination in Self-Organizing Interaction Networks

Joshua L. Payne and Jason H. Moore

Computational Genetics Laboratory, Dartmouth Medical School,
Lebanon, NH 03756, USA
Joshua.L.Payne@Dartmouth.edu

Abstract. We build on recent advances in the design of self-organizing interaction networks by introducing a sexual variant of an existing asexual, mutation-limited algorithm. Both the asexual and sexual variants are tested on benchmark optimization problems with varying levels of problem difficulty, deception, and epistasis. Specifically, we investigate algorithm performance on Massively Multimodal Deceptive Problems and NK Landscapes. In the former case, we find that sexual recombination improves solution quality for all problem instances considered; in the latter case, sexual recombination is not found to offer any significant improvement. We conclude that sexual recombination in self-organizing interaction networks may improve solution quality in problem domains with deception, and discuss directions for future research.

1 Introduction

Many natural and physical systems can be characterized as networks, where vertices denote system components and edges denote component interactions. Recent advances in computational power and the increased availability of large-scale data sets have provided several novel insights regarding the influence of network structure on the functionality and vulnerability of these systems [15], and on the dynamical processes that occur within them [4]. For example, at the cellular level, the average connectivity of a genetic regulatory network affects both its ability to discover novel phenotypes and its robustness to perturbation [3]; at the population level, the spatial localization of interaction events affects the maintenance of genetic diversity [12,22], the evolution of cooperation [16,10], and the emergence of altruistic behavior [25].

Inspired by the interaction networks of such complex adaptive systems, recent advances in cellular evolutionary algorithms have employed heterogeneous interaction networks as population structures. The saturation dynamics of advantageous alleles have been investigated in both small-world [7] and scale-free [7,19] interaction networks, and scale-free population structures have been analyzed in the context of genetic algorithms for single [6] and multiobjective optimization problems [13,14], and in evolutionary robotics applications [5].

While such complex population structures bear a closer resemblance to some natural systems than their lattice-based predecessors [24,23,8], the analogy only

goes so far. In all of the examples cited above, the interaction network was generated prior to the evolution of the population, and its structure was held fixed throughout the evolutionary process. Though this approach provides useful insights regarding the influence of various topological properties on evolutionary search [19], it is a gross oversimplification of the interaction networks of natural systems, which are both dynamic and self-organizing. To date, only two studies have attempted to incorporate these salient features of biological systems into cellular evolutionary algorithms [2,26]. In [2], the global ratio of the horizontal and vertical dimensions of a two-dimensional lattice population structure was adaptively altered via a feedback loop with the evolutionary dynamics of the population, offering significant performance improvements over static lattice-based interaction networks. However, the local neighborhood structure of the interaction network was assumed *a priori* and held constant throughout the run and the global population structure was assumed to be regular. In contrast, the algorithm presented in [26] is not only dynamic, but also allows for the self-organization of irregular local neighborhood structures and the emergence of global network properties similar to the complex adaptive systems from which nature-based optimization algorithms draw their inspiration. The results of [26] demonstrate that mutation-limited genetic algorithms structured on such self-organizing interaction networks improve both diversity maintenance and solution quality over panmictic and lattice-based population structures, in specific problem domains.

Here, we extend the algorithm presented in [26] to include sexual recombination, a variation operator that is integral to the exploration phase of evolutionary search. We compare the asexual and sexual variants of this algorithm on benchmark optimization problems with varying levels of problem difficulty, multimodality, and epistasis.

2 Methods

2.1 Self-Organizing Interaction Networks

The details of the asexual variant of the self-organizing interaction networks considered in this study are provided in [26]. For completeness, we provide a high-level overview of this algorithm here.

The interaction network is initialized as a ring of M vertices, where each vertex connects to its two nearest neighbors. A population of size M is then randomly initialized, and each individual is placed in its own vertex. Since vertices are always occupied by a single individual, the terms vertex and individual will be used interchangeably.

The coupled evolution of the population and the interaction network is broken into two phases per generation. In the first phase, M parents are selected to produce offspring, each of which is placed in a new vertex. This results in a network with $2M$ vertices. An offspring vertex connects to its parent vertex, and inherits each of its parent's edges with probability p_{add} . If an offspring inherits an edge from its parent, then the parent loses that edge with probability p_{remove} .

Since offspring connect to their parents, and a parent can only lose an edge if its offspring gains that edge, it is guaranteed that the network will remain connected (i.e., a finite path exists between all vertices). Parents are selected with uniform probability from the population, with replacement. Thus, parent selection does not depend on fitness. Offspring faithfully inherit the parental genotype, subject to mutation.

In the second phase, selection pares the interaction network back down to M vertices, as follows. A vertex is selected at random with uniform probability from the population. This vertex then competes with its lowest-fitness neighbor. Of this pair, the lower fitness vertex is removed from the network and the higher fitness vertex inherits all of its links. Since the individual with the highest fitness cannot lose one of these competitions, elitism is implicitly included in this framework.

A key insight of [26] is that the fitness of a vertex should be considered in the context of its local neighborhood. To do this, each individual is given a rank r based on how its raw fitness f compares with the raw fitnesses of its d neighbors. A rank of $r = 0$ denotes the best individual in a neighborhood and a rank of $r = d$ denotes the worst. Based on the individual's rank r , it is assigned a contextual fitness f' according to

$$f' = \frac{d - r}{d}, \quad (1)$$

which is used in all competition events. After an individual is removed from the network, the contextual fitnesses of all its neighbors are reevaluated.

These two phases are repeated for a specified number of generations. Each vertex addition and removal transforms the network away from its original homogeneous form to something more closely resembling the topologies of complex, natural systems [26] (Fig. 1).

2.2 Sexual Recombination

Here, we propose a simple extension of the asexual reproduction phase considered in [26]. When a vertex i is chosen for reproduction, a mate j is subsequently selected at random from the neighborhood of vertex i . A recombination operator is then applied to the individuals situated in vertices i and j to form an offspring in vertex z , which attaches to i and inherits i 's links in the same manner as the asexual variant.

2.3 Benchmark Problems

In this section, we briefly describe the two benchmark optimization problems used in this study. These problems were chosen because they possess important characteristics of real optimization tasks, such as multimodality, deception, and epistasis, and allow for a tunable degree of problem difficulty. Due to space constraints, we limit our description of these problems to their defining characteristics and the details needed to replicate our experiments. The interested reader should refer to [2,11] for more details.

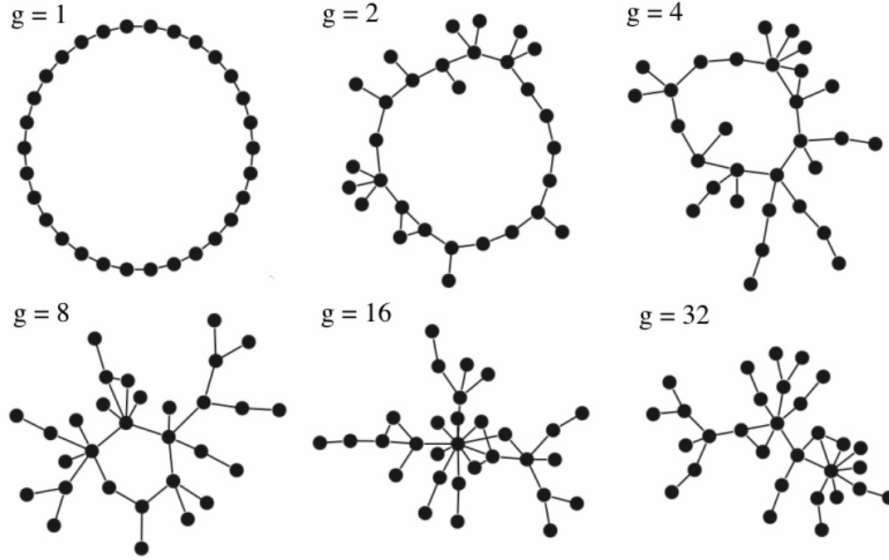


Fig. 1. Visualization of the evolutionary dynamics of self-organizing interaction networks. The network is initialized as a ring and the vertex addition and removal events of subsequent generations (g) transform the network into a heterogeneous form. For visual clarity, these networks are deliberately smaller than any used in the experiments ($M = 30$ vertices). These networks were generated by evolving a population on NK landscapes with $N = 30$ and $K = 2$ (see Section 2.3).

Massively Multimodal Deceptive Problems. Massively multimodal deceptive problems [2,9] consist of k concatenated subproblems, each with two global optima and a single deceptive suboptimum. The subproblems are six bits long, and the fitness contribution of each subproblem depends on the unification of the bits. Specifically, the unification is used as an index into the vector $(1.00, 0.00, 0.36, 0.64, 0.36, 0.00, 1.00)$. Thus, the global optima of each subproblem are at maximum hamming distance from one another (at a unification of zero and six) and provide a fitness contribution of one. The deceptive suboptimum provides a fitness contribution of 0.64 and is located at a unification of three. Maximum fitness is k , which we normalize to a maximum of 1.

NK Landscapes. NK landscapes are abstract models of fitness surfaces [11]. Each of the N bits in a binary string epistatically interact with K neighboring bits to provide a fitness contribution. These contributions are in the range $(0, 1)$ and are extracted from a randomly generated look-up table with 2^{K+1} entries for each of the N bits. The ruggedness of the fitness surface increases with the number of interacting bits K . Fitness is defined as the arithmetic mean of the N fitness contributions.

2.4 Experimental Design

To be consistent with [26], we use a population size of $M = 100$, a mutation rate of $1/N$ (where N is the length of the binary string), and $p_{add} = p_{remove} = 0.1$. We use bit-flip mutation and, in the sexual case, one-point crossover (in addition to mutation). For the MMDP, we consider $20 \leq k \leq 640$ and allow evolution to proceed for 1000 generations. (Preliminary experimentation with maximum generation time indicated that 1000 generations was sufficient for population convergence to occur, for all values of k considered.) For each problem instance, we perform 500 independent replications, where the asexual and sexual variants are seeded with the same initial populations. For the NK landscapes, we consider $N = 30$ and $2 \leq K \leq 14$, and allow the population to evolve for 5000 generations, consistent with [26]. For each value of K , we randomly generate 500 problem instances. For each problem instance, we generate a random initial population and use it to seed both algorithm variants. This experimental design allows for a paired statistical comparison of all results.

3 Results

3.1 Massively Multimodal Deceptive Problems

For all instances of the MMDP considered, the asexual and sexual algorithm variants both converged on suboptimal, deceptive solutions. However, the sexual variant always found higher fitness solutions. For example, in Fig. 2a, we depict the average best fitness of the two algorithm variants on a representative MMDP instance, with $k = 80$. In this case, the average best solution found by the asexual variant had a fitness of 0.687 (± 0.0004 s.e.), whereas the average best solution found by the sexual variant had a fitness of 0.717 (± 0.0005 s.e.).

This trend was consistent across all MMDP instances. In Fig. 3, we depict the fitness of the best solutions found at the end of each replication as a function of problem difficulty, for the asexual (A) and sexual (S) algorithm variants. The average best fitness of the sexual variant was always significantly higher than the average best fitness of its asexual counterpart ($p < 0.05$, paired t-test). As problem difficulty increased, the average best fitness found by both algorithm variants decreased (Fig. 3), an expected result given that the population size was held constant for all experiments.

3.2 NK Landscapes

In Fig. 4, we compare the fitness of the best solutions found by the asexual (A) and sexual (S) algorithms on NK landscapes with $N = 30$ and varying levels of epistasis (K). For both algorithm variants, the fitness of the best solutions varied non-monotonically with K , such that the highest fitness was observed for $6 \leq K \leq 8$. For all values of K , the distributions of best fitness were statistically indistinguishable between the two algorithms ($p > 0.05$, paired t-test).

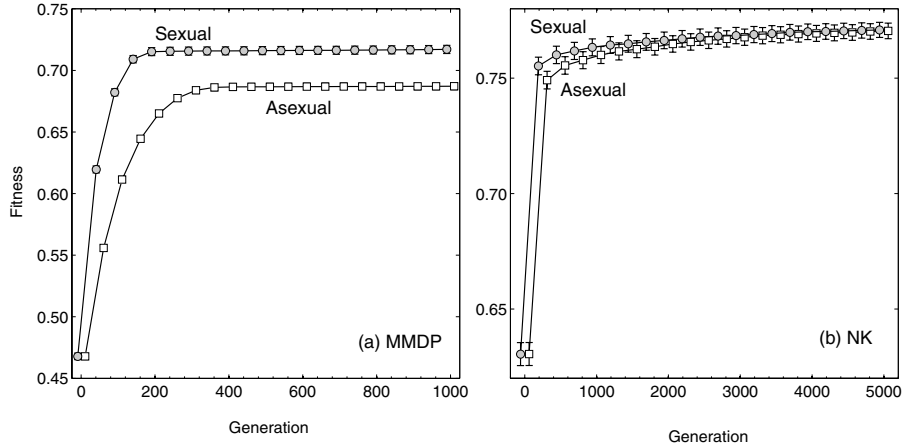


Fig. 2. Best raw fitness per generation of the sexual and asexual algorithm variants on the (a) MMDP ($k = 80$) and (b) NK ($N = 30, K = 10$) test problems. Data points are the average of all replications, and error bars denote standard error. (In (a) the error bars are smaller than the symbol size.) The data are deliberately offset in the horizontal dimension for visual clarity.

As an illustrative example, we depict in Fig. 2b the fitness of the best solutions found by the two algorithms as a function of generation number, for a representative case of intermediate epistasis ($K = 10$). On these problem instances, the best solutions found by both algorithm variants had an average fitness of 0.770 (± 0.004 s.e.).

4 Discussion

In this study, we have demonstrated that including sexual recombination in self-organizing interaction networks can improve solution quality in some problem domains. In a direct comparison between an asexual, mutation-limited algorithm [26] and a sexual, recombinative variant, we found that the sexual variant discovered higher quality solutions when the problem instance was deceptive, but offered no advantage when the problem instance was epistatic.

Our results on NK landscapes contrast those observed with panmictic genetic algorithms [1], where recombination was shown to offer an advantage over a mutation-limited variant for intermediate levels of epistasis ($12 < K < 32$ for NK landscapes with $N = 96$). We found that the asexual and sexual variants produced statistically indistinguishable results for all values of K considered. This result follows from the fact that both the ruggedness of a fitness landscape and the fitness difference between peaks and valleys increases monotonically with K [17]. Thus, not only is recombination more likely to produce offspring in maladaptive fitness valleys as K increases, but the sheer depth of these valleys acts to preclude any

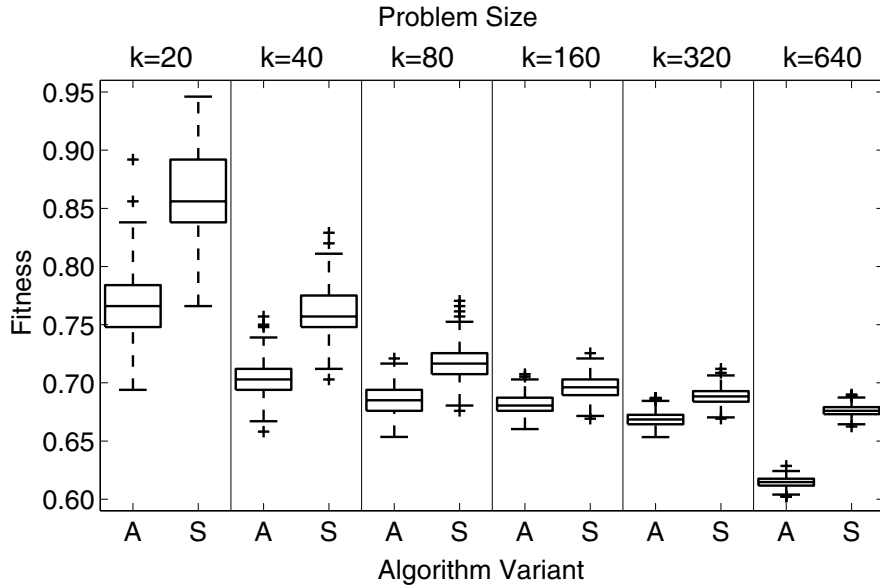


Fig. 3. Best raw fitness after 1000 generations for various MMDP problem sizes (k), using the asexual (A) and sexual (S) algorithm variants. The sexual variant produced significantly higher quality solutions in all cases ($p < 0.05$, paired t-test).

possibility of mutational escape. While such maladaptive movements in genome-space are often considered prerequisites for adaptive peak shifts in epistatic fitness landscapes [17], they were not found to offer any advantage here.

The coupling of population dynamics and population structure [26] makes these self-organizing interaction networks unique among those employed in cellular evolutionary algorithms, and allows for better diversity maintenance than static, lattice-based interaction networks [26]. An analysis of several structural properties of the interaction networks evolved in this study, including characteristic path length, clustering coefficient, degree distribution, and assortativity, did not reveal any significant difference between the asexual and sexual cases. Further, we did not find any relationship between the structure of the evolved interaction networks and the underlying problem domain or its corresponding difficulty. Future work will seek to investigate the relationship between the topological characteristics of self-organizing interaction networks and the population-level distribution of genetic information, in order to better understand how diversity is maintained in both the asexual and sexual cases. This will complement previous analyses of population distributions in static, regular interaction networks [20].

Self-organizing population structures offer the potential to improve the efficacy of evolutionary search, and better mimic some of the features of complex biological systems. However, several important aspects of natural interaction

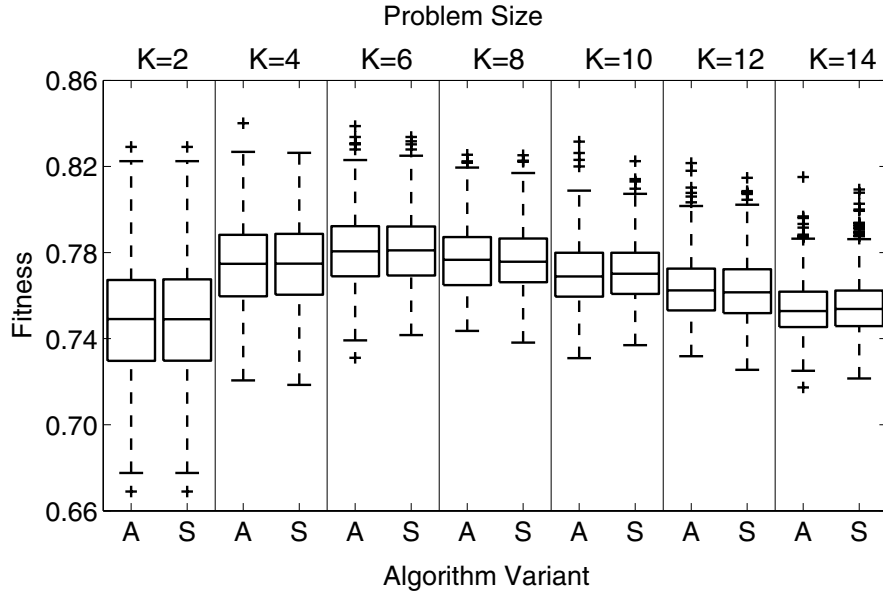


Fig. 4. Best raw fitness after 5000 generations as a function of the degree of epistasis K in NK landscapes with $N = 30$ for the asexual (A) and sexual (S) algorithm variants. Solution quality was statistically indistinguishable between the asexual and sexual algorithm variants ($p > 0.05$, paired t-test).

networks are not included in this framework. In biotic populations, individuals often choose their own social ties, as opposed to inheriting them from their parents, as was the case here. This ‘active linking’ has offered novel insights into such pertinent topics as the evolution of cooperation, in both theoretical [18] and experimental settings [21]. We believe that allowing individuals to strengthen (or weaken) their ties with competitors or mating partners may further improve the search capabilities of cellular evolutionary algorithms that employ self-organizing interaction networks. Current research is directed along these lines.

References

1. Aguirre, H.E., Tanaka, K.: Genetic algorithms on NK-landscapes: Effects of selection, drift, mutation, and recombination. In: Raidl, G.R., Cagnoni, S., Cardalda, J.J.R., Corne, D.W., Gottlieb, J., Guillot, A., Hart, E., Johnson, C.G., Marchiori, E., Meyer, J.-A., Middendorf, M. (eds.) *EvoIASP 2003, EvoWorkshops 2003, EvoSTIM 2003, EvoROB/EvoRobot 2003, EvoCOP 2003, EvoBIO 2003, and EvoMUSART 2003*. LNCS, vol. 2611, pp. 131–142. Springer, Heidelberg (2003)
2. Alba, E., Dorronsoro, B.: The exploration/exploitation tradeoff in dynamic cellular genetic algorithms. *IEEE Transactions on Evolutionary Computation* 9(2), 126–142 (2005)

3. Aldana, M., Balleza, E., Kauffman, S.A., Resendis, O.: Robustness and evolvability in genetic regulatory networks. *Journal of Theoretical Biology* 245, 433–448 (2007)
4. Barrat, A., Barthélemy, M., Vespignani, A.: *Dynamical Processes on Complex Networks*. Cambridge University Press, Cambridge (2008)
5. Gasparri, A., Panzieri, S., Pascucci, F., Ulivi, G.: A spatially structured genetic algorithm over complex networks for mobile robot localisation. In: *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 4277–4282. IEEE Press, Los Alamitos (2007)
6. Giacobini, M., Preuss, M., Tomassini, M.: Effects of scale-free and small-world topologies on binary coded self-adaptive cea. In: Gottlieb, J., Raidl, G.R. (eds.) *Evolutionary Computation and Combinatorial Optimization*, pp. 86–98. Springer, Heidelberg (2006)
7. Giacobini, M., Tomassini, M., Tettamanzi, A.: Takeover times curves in random and small-world structured populations. In: Beyer, H.G. (ed.) *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2005*, pp. 1333–1340. ACM Press, New York (2005)
8. Giacobini, M., Tomassini, M., Tettamanzi, A., Alba, E.: Selection intensity in cellular evolutionary algorithms for regular lattices. *IEEE Transactions on Evolutionary Computation* 9(5), 489–505 (2005)
9. Goldberg, D., Deb, K., Horn, J.: Massive multimodality, deception, and genetic algorithms. In: Männer, R., Manderick, B. (eds.) *Parallel Problem Solving From Nature*, pp. 37–46. North-Holland, Amsterdam (1992)
10. Hauert, C., Doebeli, M.: Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* 428, 643–646 (2004)
11. Kauffman, S.A.: *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, Oxford (1993)
12. Kerr, B., Riley, M.A., Feldman, M.W., Bohannan, B.J.M.: Local dispersal promotes biodiversity in a real life game of rock-paper-scissors. *Nature* 418, 171–174 (2002)
13. Kirley, M., Stewart, R.: An analysis of the effects of population structure on scalable multiobjective optimization problems. In: Thierens, D. (ed.) *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2007*, pp. 845–852. ACM Press, New York (2007)
14. Kirley, M., Stewart, R.: Multiobjective optimization on complex networks. In: Obayashi, S., Deb, K., Poloni, C., Hiroyasu, T., Murata, T. (eds.) *EMO 2007. LNCS*, vol. 4403, pp. 81–95. Springer, Heidelberg (2007)
15. Newman, M.E.J., Barabási, A.L., Watts, D.J. (eds.): *The Structure and Dynamics of Networks*. Princeton University Press, Princeton (2006)
16. Nowak, M.A., May, R.M.: Evolutionary games and spatial chaos. *Nature* 359, 826–829 (1992)
17. Ostman, B., Hintze, A., Adami, C.: Impact of epistasis on evolutionary adaptation. *arXiv:0909.3506v1* (2009)
18. Pacheco, J.M., Traulsen, A., Ohtsuki, H., Nowak, M.A.: Repeated games and direct reciprocity under active linking. *Journal of Theoretical Biology* 250, 723–731 (2008)
19. Payne, J.L., Eppstein, M.J.: Evolutionary dynamics on scale-free interaction networks. *IEEE Transactions on Evolutionary Computation* 13(4), 895–912 (2009)
20. Preuss, M., Lasarczyk, C.: On the importance of information speed in structured populations. In: Yao, X., Burke, E.K., Lozano, J.A., Smith, J., Merelo-Guervós, J.J., Bullinaria, J.A., Rowe, J.E., Tiño, P., Kabán, A., Schwefel, H.-P. (eds.) *PPSN 2004. LNCS*, vol. 3242, pp. 91–100. Springer, Heidelberg (2004)
21. Rand, D.R., Dreber, A., Ellingsen, T., Fudenberg, D., Nowak, M.A.: Positive interactions promote public cooperation. *Science* 325, 1272–1275 (2009)

22. Reichenbach, T., Mobilia, M., Frey, E.: Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature* 448, 1046–1049 (2007)
23. Rudolph, G.: On takeover times in spatially structured populations: array and ring. In: Lai, K.K., Katai, O., Gen, M., Lin, B. (eds.) *Proceedings of the Second Asia-Pacific Conference on Genetic Algorithms and Applications, APGA-2000*, pp. 144–151. Global Link Publishing Company, Hong Kong (2000)
24. Sarma, J., De Jong, K.: An analysis of the effect of neighborhood size and shape on local selection algorithms. In: Ebeling, W., Rechenberg, I., Voigt, H.-M., Schwefel, H.-P. (eds.) *PPSN 1996. LNCS*, vol. 1141, pp. 236–244. Springer, Heidelberg (1996)
25. Werfel, J., Bar-Yam, Y.: The evolution of reproductive restraint through social communication. *Proceedings of the National Academy of Science* 101(30), 11019–11024 (2004)
26. Whitacre, J.M., Sarker, R.A., Pham, Q.T.: The self-organization of interaction networks for nature-inspired optimization. *IEEE Transactions on Evolutionary Computation* 12(2), 220–230 (2008)