Ecological Complexity 30 (2017) 47–56

Contents lists available at ScienceDirect

Ecological Complexity

journal homepage: www.elsevier.com/locate/ecocom

Original Research Article

Effect of group organization on the performance of cooperative processes

Sandro M. Reia, José F. Fontanari *

Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369, 13560-970 São Carlos, São Paulo, Brazil

ARTICLE INFO

Article history:
Received 28 April 2016
Received in revised form 3 September 2016
Accepted 12 September 2016
Available online 20 September 2016

Keywords:
Imitative learning
Group organization
Collective brain
Social networks

ABSTRACT

Problem-solving competence at group level is influenced by the structure of the social networks and so it may shed light on the organization patterns of gregarious animals. Here we use an agent-based model to investigate whether the ubiquity of hierarchical networks in nature could be explained as the result of a selection pressure favoring problem-solving efficiency. The task of the agents is to find the global maxima of NK fitness landscapes and the agents cooperate by broadcasting messages informing on their fitness to the group. This information is then used to imitate, with a certain probability, the fittest agent in their influence networks. The performance of the group is measured by the time required to find the global maximum. For rugged landscapes, we find that the modular organization of the hierarchical network with its high degree of clustering eases the escape from the local maxima, resulting in a superior performance as compared with the scale-free and the random networks. The optimal performance in a rugged landscape is achieved by letting the main hub to be only slightly more propense to imitate the other agents than vice versa. The performance is greatly harmed when the main hub carries out the search independently of the rest of the group as well as when it compulsively imitates the other agents.

1. Introduction

There is little dispute over the claim that the collective structures built by termites, ants and slime molds are products of cooperative work performed by a myriad of organisms who, individually, are inept to conceive the greatness of the structures they build (Marais, 1937). Thinking of those collective structures as the organisms' solutions to the problems that endanger their existence, it is natural to argue that competence in problem solving should be viewed as a candidate selection pressure for molding the organization of groups of social animals (Bloom, 2001; Queller and Strassmann, 2009).

Information flows between individuals via social contacts and, in the problem-solving context, the relevant process is imitative learning as expressed in this quote by Bloom "Imitative learning acts like a synapse, allowing information to leap the gap from one creature to another" which summarizes his view of those collective structures as global brains (Bloom, 2001). Evidences that cooperative work powered by social learning is an efficient process to solve difficult problems are offered by the variety of social learning based optimization heuristics, such as the particle swarm optimization algorithm (Bonabeau et al., 1999) and the adaptive culture heuristic (Kennedy, 1998; Fontanari, 2010).

From the perspective of the computer science, there has been considerable progress on the understanding of the factors that make cooperative group work effective (Clearwater et al., 1991, 1992; Page, 2007), although, somewhat disturbingly, the most popular account of collective intelligence, the so-called wisdom of crowds, involves the suppression of cooperation since its success depends on the individuals making their guesses independently of each other (Surowiecki, 2004) (see, however, King et al., 2012). We note that quite recently formalized approaches to collective learning dynamics have been considered by the applied mathematics community (Bonacich and Lu, 2012; Burini et al., 2016).

In this contribution we build on a recently proposed minimal model of distributed cooperative problem-solving systems based on imitative learning (Fontanari, 2014) to study the influence of the social network topology on the performance of cooperative processes. Individuals cooperate by broadcasting messages informing on their fitness and use this information to imitate, with a certain probability, the fittest individual in their influence networks. The task of the individuals is to find the global maxima of smooth and rugged fitness landscapes generated by the NK model

* Corresponding author.
E-mail addresses: smreia@gmail.com (S.M. Reia), fontanari@ifsc.usp.br (J.F. Fontanari).

http://dx.doi.org/10.1016/j.ecocom.2016.09.002
1476-945X/

(Kauffman and Levin, 1987; Kauffman, 1989) and the performance or efficiency of the group is measured by the number of trials required to find those maxima.

Our goal is to investigate whether the ubiquity of hierarchical networks (Ravasz and Barabási, 2003), which are both modular and scale-free, could be explained as the result of a selection pressure favoring problem-solving efficiency. In fact, for rugged landscapes we find that the hierarchical network performs better than the (non-modular) scale-free and the random networks. The modular organization of the hierarchical network with its high degree of clustering facilitates the system to escape local maxima, despite the presence of a hub with very high connectivity (super-spreader) which, in general, may cause great harm to the system performance by broadcasting misleading information about the location of the global maxima (Fontanari and Rodrigues, 2016) as happens in the case of the (non-modular) scale-free network. For smooth landscapes, the topology of the network has little influence on the performance of the imitative search. Since the hierarchical networks are a very small subclass of the scale-free networks, henceforth we will use the expression scale free network to refer to a non-modular scale-free network and the expression hierarchical network to refer to a modular scale-free network.

In addition, we find that for the three network topologies considered here, namely, hierarchical, scale-free and random topologies, allowing the main hub (i.e., the node with the highest degree) to explore the landscape without much consideration for the other individuals, even though those individuals may learn from it, is always detrimental to the performance of the system. Interestingly, for the hierarchical and scale-free networks, the optimal performance in a rugged landscape is achieved by letting the main hub to be only slightly more propense to imitate the other agents than vice versa. A compulsive imitator located at the main hub of the hierarchical network (or at the two main hubs of a scale-free network) leads to a disastrous performance. For the random network, where the main hub is not very influential, the performance is maximized by the compulsive imitation strategy. This is also true for the three topologies in the case of a smooth landscape, but the reason is that in the absence of local maxima it is always better to imitate the fittest individual in the group.

In the cooperative problem-solving context, cooperation means the exchange of information between agents that may, in principle, allow them to find the solution to a common problem more rapidly than if they worked in isolation. In that scenario, the opposite of cooperation is independent work and there is no conflict of interests between the agents in the group. This usage of the word cooperation contrasts with the meaning of cooperation in evolutionary game theory, where there is a conflict of interests between the agents and, consequently, a cost for the cooperative agents (Axelrod, 1984).

The rest of this paper is organized as follows. For the sake of completeness, we present a brief description of the NK model of rugged fitness landscapes in Section 2. The rules of the agent-based model that implements the imitative search are explained in Section 3 and the three network topologies – hierarchical, scale-free and random – are presented in Section 4. In Section 5 we present and discuss the results of the simulations of the imitative search on rugged and smooth NK landscapes for those three topologies. Finally, Section 6 is reserved to our concluding remarks.

2. NK model of rugged fitness landscapes

The NK model is a computational framework to generate families of statistically identical rugged fitness landscapes. It was proposed by Stuart Kauffman in the late 1980s aiming at modeling evolution as an incremental process, the so-called adaptive walk, on rugged landscapes (Kauffman and Levin, 1987; Kauffman, 1989). Today the NK model is the paradigm of problem spaces with many local optima, being particularly popular among the organizational and management research community (Levinthal, 1997; Lazer and Friedman, 2007; Fontanari, 2016).

The NK model is named for the two integer parameters that are used to randomly generate landscapes, namely, N and K. The landscape is defined in the space of binary strings of length N and so this parameter determines the size of the solution space, 2^N. The other parameter $K = 0, \ldots, N-1$ influences the ruggedness of the landscape. In particular, the correlation between the fitness of any two neighboring strings (i.e., strings that differ at a single component) is $1-(K+1)/N$ (Kauffman, 1989). Hence $K = 0$ corresponds to a smooth landscape whereas $K = N-1$ corresponds to a completely uncorrelated landscape. For concreteness, next we describe briefly the procedure to generate a random realization of a NK landscape.

The 2^N distinct binary strings of length N are denoted by $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ with $x_i = 0, 1$. To each string $\mathbf{x}$ we associate a fitness value $\Phi(\mathbf{x})$ which is an average of the contributions from each component i in the string, i.e., $\Phi(\mathbf{x}) = \sum_{i=1}^{N} \phi_i(\mathbf{x})/N$, where ϕ_i is the contribution of component i to the fitness of string $\mathbf{x}$. It is assumed that ϕ_i depends on the state x_i as well as on the states of the K right neighbors of i, i.e., $\phi_i = \phi_i(x_i, x_{i+1}, \ldots, x_{i+K})$ with the arithmetic in the subscripts done modulo N, i.e., the numbers wrap around upon reaching N (for example, $i + N = i$ when the sum is done modulo N). The functions ϕ_i are N distinct real-valued functions on $\{0, 1\}^{K+1}$ but the usual procedure is to assign to each ϕ_i a uniformly distributed random number in the unit interval (Kauffman, 1989), which then guarantees that $\Phi \in (0, 1)$ has a unique global maximum.

For $K = 0$ the global maximum is the sole maximum of Φ, which can be easily found by picking for each component i the state $x_i = 0$ if $\phi_i(0) > \phi_i(1)$ or the state $x_i = 1$, otherwise. For $K = N-1$, the (uncorrelated) landscape has on the average $2^N/(N+1)$ maxima with respect to single bit flips (Derrida, 1981). Finding the global maximum of the NK model for $K > 0$ is a NP-complete problem (Solow et al., 2000), which means that the time required to solve the problem using any currently known deterministic algorithm increases exponentially fast with the length N of the strings (Garey and Johnson, 1979).

We note that the specific features of a realization of the NK landscape (e.g., number and location of the local maxima with respect to the global maximum) are not fixed by the parameters N and K, because the components ϕ_i are chosen randomly in the unit interval. This is the reason that finding the global maximum for *any* realization of the NK landscape for large N and $K > 0$ is an extremely difficult computational problem (Solow et al., 2000). Hence, in order to better apprehend the influence of the network topology and, in particular, the role of the main hub on the performance of cooperative problem-solving systems, here we use a single realization of the NK fitness landscape for fixed values of N and K.

More pointedly, we consider two types of landscape: a smooth landscape with $N = 16$ and $K = 0$ and a rugged landscape with $N = 16$ and $K = 5$. Since for $K = 0$ all NK landscapes are equivalent, there is no lack of generality in considering a single instance of that family. The particular realization of the NK landscape with $N = 16$ and $K = 5$ considered here exhibits 296 maxima in total, among which 295 are local maxima. The mean relative fitness of the local maxima with respect to the fitness of the global maximum is 0.81 whereas the mean relative fitness of all strings is 0.60. It is interesting to note that for large N the NK model exhibits the so-called complexity catastrophe (Kauffman, 1989), i.e., as N increases the fitness of the local maxima become poorer to such a point that they are not better than the fitness of a randomly chosen string. The effects of averaging over different realizations of the rugged landscape is addressed briefly at the end of Section 5.