Abstract—In cooperative multiagent systems, agents interact to solve tasks. Global dynamics of multiagent teams result from local agent interactions, and are complex and difficult to predict. Evolutionary computation has proven a promising approach to design of such teams. The majority of current studies use teams composed of agents with identical control rules (“genetically homogeneous teams”) and select behavior at the team level (“team-level selection”). Here we extend current approaches to include four combinations of genetic team composition and level of selection. We compare the performance of genetically homogeneous teams evolved with individual-level selection, genetically homogeneous teams evolved with team-level selection, genetically heterogeneous teams evolved with individual-level selection, and genetically heterogeneous teams evolved with team-level selection. We use a simulated foraging task to show that the optimal combination depends on the amount of cooperation required by the task. Accordingly, we distinguish between three types of cooperative tasks and suggest guidelines for the optimal choice of genetic team composition and level of selection.

Index Terms—Altruism, artificial evolution, cooperation, evolutionary robotics, fitness allocation, multiagent systems (MAS), team composition.

I. INTRODUCTION

MULTIAGENT SYSTEMS (MAS) span a large number of research fields, from software agents to robotics, and play a key role in several industrial applications, such as ground and air vehicle control, supply chains, or network routing. The design of control rules for MAS is challenging because agent behavior depends not only on interactions with the environment but also on the behavior of other agents. As the number of interacting agents in a team grows, or when agent behaviors become more sophisticated, the design of suitable control rules rapidly becomes very complex. This is especially true when agents are expected to coordinate or cooperate to collectively achieve a desired task. Evolutionary computation has been advocated as an effective and promising strategy to generate control parameters and decision rules for collective agents [1], [2].

Manuscript received July 14, 2007; revised November 21, 2007 and June 26, 2008; accepted November 5, 2008. Current version published June 10, 2009. This work was sponsored by the Swiss National Science Foundation and the European Commission (FP6, IST-FET projects ECAgents and Swarmoids).

M. Waibel and D. Floreano are with the Laboratory of Intelligent Systems, School of Engineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland (e-mail: markus.waibel@epfl.ch; dario.floreano@epfl.ch).

L. Keller is with the Department of Ecology and Evolution, Biophore, University of Lausanne, CH-1015 Lausanne, Switzerland (e-mail: laurent.keller@unil.ch).

Digital Object Identifier 10.1109/TEVC.2008.2011741

In addition to the methodological issues of evolving agents that operate in isolation [2], the evolution of agent teams must address two major issues: 1) It must determine optimal team composition. Agents of a team may either share control rules (genetically homogeneous teams) or employ different ones (genetically heterogeneous teams), and 2) It requires a suitable method for selective reproduction of desired team behavior. Selection may operate either on individuals (individual-level selection) or on teams (team-level selection). In the simplest case, one must decide between genetically homogeneous or
heterogeneous teams, and between selecting agents at the individual or at the team level.

Fig. 1 shows a sample of previous work on the evolution of MAS in robotics, combinatorial optimization, cellular automata, artificial life, genetic programming, and others, plotted according to the chosen genetic team composition and level of selection. In addition to work cited in Fig. 1, some authors have used cooperative co-evolutionary algorithms (CCEAs, [57]) to evolve heterogeneous control rules for teams of agents [58]–[61]. CCEAs are applied by decomposing problem representations into subcomponents and then creating a separate population of individuals for each subcomponent. This approach allows teams to be composed of specialized subgroups and corresponds to the biological co-evolution of multiple species. In their basic form, CCEAs require the designer to manually decompose the multiagent task, and thus to solve part of the optimization problem beforehand. Work that used machine learning techniques other than evolutionary computation (e.g., reinforcement learning) is not considered in this paper.

Fig. 1 suggests that the majority of current approaches to the evolution of MAS use genetically homogeneous teams evolved with team-level selection. Where the reasons for the choice of genetically homogeneous teams are made explicit, it is argued that homogeneous teams are easy to use [8], [36], require fewer evaluations [25], [32], scale more easily [13], and are more robust against the failure of team members [13], [62] than heterogeneous teams. Many other approaches use genetically heterogeneous teams evolved with individual-level selection. Genetically heterogeneous teams are sometimes seen as providing more behavioral flexibility [25] and as providing advantages in tasks that require specialization [7], [25], [62], [63].

The terms “homogeneous team” and “heterogeneous team” used in the current literature cover many different aspects. It is important to note that while all agents in genetically homogeneous teams share the same genes, agents can nevertheless be behaviorally heterogeneous. This can happen when agents differentiate during their lifetime, for example due to varying initial conditions [30], or due to developmental processes or learning [64]. This can also happen when agents “activate” different parts of their genome, for example when each agent’s behavior is controlled by a different section of a single team genome [11], [22], [28], [33]. In this case, agents can specialize on different functions, yet be genetically identical, just like specialized cells in a biological organism. Conversely, it is important to note that genetically heterogeneous teams are those in which agents are, on average, not genetically more similar to team members than to agents in the rest of the population [65], [66]. This means that teams resulting from embodied evolution or common versions of steady-state evolution are usually genetically heterogeneous although these algorithms often generate multiple offspring from a single parent, resulting in genetically similar (but not identical) agents. In some cases, teams consist of clonal sub-teams [25], [26] or agents that share only part of their genome. Teams with agents that are, on average, genetically more similar (but not identical) to members of their team than to members of the rest of the population are termed “partially heterogeneous.” The effects of partial genetic heterogeneity on the evolution of multiagent teams are not yet fully explored in evolutionary computation [39], but there is evidence that they can lead to improved specialization [25], [26]. These effects have been deeply studied in biology [67], [68].

The choice of level of selection is rarely discussed explicitly. Some research has addressed the related issue of credit assignment for the evolution of MAS [40], [69]. In the context of MAS, credit assignment is concerned with distributing fitness rewards among individual agents. Fitness distribution leads to credit assignment problems [70], [71] in many cooperative multiagent tasks, because individual contributions to team performance are often difficult to estimate or difficult to monitor [72]. Selection is usually performed on the basis of accumulated individual or team fitness, which may be the result of many fitness rewards with different types of credit assignment. Therefore an optimal choice of level of selection is not only influenced by the type of task but also by the types of credit assignment used.

Genetic team composition and level of selection have long been identified as two important factors for the evolution of biological multiagent teams such as groups of genes, cells, individuals, or other replicators [67], [73]. In particular the evolution of altruism [74], in which agents cooperate to increase team fitness in spite of an individual fitness cost to the cooperator, has received a lot of attention [68], [75]. Here we define cooperation as a behavior that increases the fitness of other agents, and altruistic cooperation (altruism) as a behavior that increases the fitness of other agents and decreases the cooperator’s fitness.

In this paper, we focus on cooperative multiagent tasks that do not require specialization. We compare the performance of robot teams evolved in four evolutionary conditions: genetically homogeneous teams evolved with team-level selection; genetically homogeneous teams evolved with individual-level selection; genetically heterogeneous teams evolved with team-level selection; and genetically heterogeneous teams evolved with individual-level selection. We evaluate the performance of robot teams evolved in these four evolutionary conditions for three classes of multirobot tasks: a task that does not require cooperation; a task that requires cooperation but does not imply a cost for cooperators; and a task that requires altruistic cooperation, i.e., a task that implies an individual fitness cost for cooperators. Cooperative tasks that benefit from specialization were not considered in this paper.

II. Evolutionary Conditions

The four possible combinations of genetic team composition and level of selection were formalized into four evolutionary algorithms (Fig. 2). For the remainder of the paper we will use the terms “homogeneous” and “heterogeneous” to designate genetically homogeneous and genetically heterogeneous teams, respectively, and the terms “individual selection” and “team selection” to designate teams evolved with individual-level selection and team-level selection, respectively. We consider populations composed of $M$ teams, each composed of $N$
Algorithm 1 Homogeneous teams, individual selection

for each of \( M \) new teams do
    select two individuals from all old teams
    recombine their genomes to create one new genome
    mutate new genome
    clone new genome to obtain \( N \) genomes for new team
end for

Algorithm 2 Homogeneous teams, team selection

for each of \( M \) new teams do
    select two old teams
    recombine their genomes to create one new genome
    mutate new genome
    clone new genome to obtain \( N \) genomes for new team
end for

Algorithm 3 Heterogeneous teams, individual selection

for each of \( M \) new teams do
    for each of \( N \) new team members do
        select two individuals from all old teams
        recombine their genomes to create one new genome
        mutate new genome
        add new genome to new team
    end for
end for

Evolutionary algorithms were thus genetically homogeneous. For a new generation, each of the \( M \) new teams was created from two old teams selected using roulette-wheel selection. The two genomes of the selected teams were recombined (one-point crossover, crossover probability of 0.05) to produce one new genome. The resulting new genome was mutated by flipping the value of each bit with a probability of 0.05 and then cloned \( N - 1 \) times to obtain the \( N \) robot genomes of the new team. Teams evolved using this evolutionary condition were thus genetically homogeneous.

1) Algorithm 1—Homogeneous Teams, Individual Selection: Each of the \( M \) teams at generation 0 was formed by generating one random genome and cloning it \( N - 1 \) times to obtain \( N \) identical robot genomes (clones) per team. Teams were evaluated in the task and an individual fitness determined for each of the \( N \) robots. For a new generation, each of the \( M \) new teams was created from two \( M \) teams at generation 0. Individuals' genomes were binary strings.

2) Algorithm 2—Homogeneous Teams, Team Selection: Each of the \( M \) teams at generation 0 was formed by generating one random genome and cloning it \( N - 1 \) times to obtain \( N \) identical robot genomes (clones) per team. Teams were evaluated in the task, and for each team a team fitness was determined as the sum of the individual fitnesses of all \( N \) robots. For a new generation, each of the \( M \) new teams was formed by generating \( N \) random genomes. Teams were evaluated in the task and an individual fitness determined for each of the \( N \) robots. For a new generation, each of the \( N \times M \) new individuals was created from two \( M \) teams at generation 0. Teams evolved using this evolutionary condition were thus genetically homogeneous.

3) Algorithm 3—Heterogeneous Teams, Individual Selection: Each of the \( M \) teams at generation 0 was formed by generating \( N \) random genomes. Teams were evaluated in the task and an individual fitness determined for each of the \( N \) robots. For a new generation, each of the \( N \times M \) new individuals was created from two \( M \) teams at generation 0. Teams evolved using this evolutionary condition were thus genetically heterogeneous.

4) Algorithm 4—Heterogeneous Teams, Team Selection: Each of the \( M \) teams at generation 0 was formed by generating \( N \) random genomes. Teams were evaluated in the task, and for each team a team fitness was determined as the sum of the individual fitnesses of all \( N \) robots. For a new generation, each of the \( N \times M \) new individuals was created from two \( M \) teams at generation 0. Teams evolved using this evolutionary condition were thus genetically heterogeneous.
III. EXPERIMENTAL METHOD

A. Scenario

The experimental setup (Fig. 3) consisted of a 50 × 50 cm² arena with 10 micro-robots and two types of tokens: small and large. We chose to study a foraging task, because foraging combines several aspects of multiagent tasks (distributed search, coordinated movement, transportation) and relates to many real-world problems [76], [77]. In addition, foraging is a widespread and well-studied behavior of many biological societies [78]–[80]. Experiments were conducted in simulation of micro-robots and evolved controllers were transferred to the real robots (see Section III-D).

Robots foraged tokens by transporting them into a 4-cm wide region at one side of the arena marked by a white wall. A single robot was sufficient to transport a small token. At least two robots were required to transport a large token, and thus retrieval of large tokens required cooperation. Cooperating agents had to coordinate their behaviors to successfully align their positions before and during token transport.

The micro-robots [81] were small (2 × 2 × 4 cm³) two-wheeled robots equipped with three infrared distance sensors at the front and one at the back, which could sense objects up to 3 cm away and allowed robots to distinguish between small and large tokens (Fig. 4 left). An extension module with a fourth infrared distance sensor with a range of up to 6 cm and a linear camera were mounted higher on the robot, overlooking tokens but sensitive to other robots and walls.

B. Control and Genetic Architecture

Robots were controlled by a feed-forward neural network with a single layer of three hidden neurons (Fig. 4 right) and a sigmoid activation function (tanh). The inputs were given by the activation values of five infrared sensors, two vision sensors, and a constant bias value of −1. Infrared sensor activation values were scaled in the range [0; 1]. Vision sensors were an average of three equidistantly spread camera pixels spanning a field of view of 18°, for the left or right side of the image, respectively. The averages were thresholded to yield 0 for a white or 1 for a black arena wall. Using the average value of three pixels rather than a single pixel allowed a robust detection of the white foraging target area in spite of the presence of other robots in the field of view. The two output units were used to control the left and right wheel motors. The activation values in the range [−1; 1] were mapped into speeds in the range [−4; 4] cm/s, with speeds in the interval of [−2.5; 2.5] cm/s set to 0 because of unreliable motor response at low speeds.

The neural network connection weights were in the range of [−2; 2] and coded on 8 bits. The genome of one individual was thus 8 × 32 bits long.
neural network with a single layer of three hidden neurons. Inputs were given by the activation values of five infrared (IR) sensors and two vision sensors with activation values computed from left and right camera pixels (see text).

This allowed robots to distinguish tokens from walls and other robots. Right: The neural network architecture, which is a feed-forward neural network with a single layer of three hidden neurons.

C. Collective Tasks

We devised three types of foraging tasks that differed in the amount of cooperation required from agents.

1) Task 1—Individual Foraging: The arena contained 6 small tokens, each of which awarded 1 fitness point to the foraging robot. This task did not require cooperation, because a single agent was sufficient to transport a small token.

2) Task 2—Cooperative Foraging: The arena contained four large tokens, which each awarded 1 fitness point to each team member, irrespective of its participation in the token foraging. This corresponded to a situation where the individual contributions to team performance were not known, i.e., a situation with credit assignment problems [70], [71], which is the case for many cooperative multiagent tasks [72]. This task required cooperation because it could not be accomplished by a single agent.

3) Task 3—Altruistic Cooperative Foraging: The arena contained six small and four large tokens. Small tokens each awarded one fitness point to the foraging robot and large tokens each awarded one fitness point to each team member, irrespective of their participation in the token foraging. In this task cooperation was costly for individuals, because individuals that did not cooperate always had higher fitness than their cooperating team mates. This meant that cooperators suffered a relative individual fitness cost and therefore this task required altruistic cooperation [68].

D. Evolutionary Experiments

Due to the large number of evaluations required for the evolution of robot behaviors, all evolutionary experiments were conducted using a physics-based 2-D simulator [82], which is available as part of an open evolutionary framework [83]. All simulation parameters, including robot size, shape, speed, and weight, as well as collision dynamics, friction forces, and sensor and actuator modalities, were based on the micro-robots described in Section III-A.

We evolved teams of robots under the four evolutionary conditions separately for each of the three tasks, making a total of 12 experimental lines. Evolutionary experiments lasted for 300 generations. Twenty independent runs were performed for each experimental line. Populations consisted of 100 teams of 10 agents each. Each team was evaluated 10 times for 3 minutes with random token and robot starting positions and orientations. Fitness was averaged over the 10 evaluations.

To compare the efficiency of the four evolutionary conditions, we re-evaluated the best teams at generation 300 for 1000 times and compared their team fitness. Fitness values were analyzed using Wilcoxon rank sum tests. All fitness values were normalized for each task, with 0 being the minimal possible fitness and 1 the theoretical maximum value.

IV. RESULTS

A. Task 1—Individual Foraging

Successful foraging behavior evolved for all four evolutionary conditions (Fig. 5). After 300 generations of artificial evolution, heterogeneous teams evolved with individual selection collected all 10 tokens in most evaluations and achieved fitness values close to the maximum value achievable. These fitness values were higher than those of homogeneous teams evolved with individual selection and homogeneous teams evolved with team selection (Wilcoxon rank sum test, $df = 38$, \( P < 0.001 \) and \( P < 0.006 \), respectively). A possible reason could be the disparities in genome evaluation in homogeneous and heterogeneous teams. For a team size of \( N \) agents, heterogeneous teams evaluated \( N \) times more genomes than homogeneous teams. This was because each heterogeneous team consisted of \( N \) different genomes, whereas homogeneous teams consisted of \( N \) identical genomes. On the other hand, homogeneous teams evaluated each genome \( N \) times more.
often than heterogeneous teams. This was because each team evaluation evaluated an identical genome $N$ times. Our results suggest that higher evaluation accuracy may have been less important than a larger number of different genomes in this task. The larger number of genomes may have allowed heterogeneous teams to discover solutions faster than homogeneous teams, which could explain the steep initial fitness increase. It may also have allowed heterogeneous teams to discover better solutions than homogeneous teams, which could explain the higher final fitness obtained with this evolutionary condition. To test whether these disparities in genome evaluation caused the high team performance of heterogeneous teams evolved with individual selection, we performed a set of additional experiments (see additional experiments without these disparities in the next section).

Performance of homogeneous teams evolved with individual selection and homogeneous teams evolved with team selection did not differ significantly ($P = 0.337$). This was because with roulette-wheel selection the probability of a team to be selected was the same as the sum of the probabilities of each individual team member to be selected. Since all team members of homogeneous teams shared the same genome, selection probabilities for a given genome were equal for both homogeneous evolutionary conditions. It should be noted, however, that this is not necessarily true for other types of selection. Selection mechanisms where the fitness of a genome is not directly proportional to its probability to be selected (e.g., truncation or rank-based selection) may lead to differences in the number of selected individuals with a given genotype and consequently affect the relative performance of homogeneous teams evolved with individual selection and homogeneous teams evolved with team selection. In these cases, individual selection may select for genomes that lead to higher maximum but lower average individual performance. However, additional investigations using truncation selection (selection of best 30% of the population; all other experimental conditions identical) did not find such performance differences in any of the three types of task ($P = 0.350/0.903/0.394$ for tasks 1/2/3 respectively; see Fig. S4 in the supplementary material).

Heterogeneous teams evolved with team selection performed significantly worse than all other evolutionary conditions.
of good genomes could only happen indirectly, by selecting those teams that contained a better mix of genomes than other teams. Since good genomes could be part of bad teams or bad genomes part of good teams, selection for good individual genomes was inefficient. This explains the slow initial fitness increase and the lowest final fitness of heterogeneous teams evolved with team selection.

B. Task 1—Individual Foraging: Disparities in Genome Evaluation and Credit Assignment

To test the hypothesis that the high team performance of heterogeneous teams evolved with individual selection was caused by disparities in genome evaluation, we performed a set of additional experiments. First, we evolved homogeneous teams in the same task, but used only one evaluation per team rather than 10 evaluations [Fig. 6(a)]. Second, we evolved heterogeneous teams in the same task, but used only 100 agents per population rather than 1000 agents [Fig. 6(b)]. In this set of experiments homogeneous and heterogeneous teams therefore evaluated the same number of genomes and had the same number of evaluations per genome.

Without disparities in genome evaluation, heterogeneous teams evolved with individual selection performed similar to homogeneous teams evolved with individual selection and homogeneous teams evolved with team selection (all three $P > 0.597$). Heterogeneous teams evolved with team selection performed worse than all other evolutionary conditions (all three $P < 0.001$), because the efficiency of selection was not affected by the changes in genome evaluation.

C. Task 2—Cooperative Foraging

Successful foraging behavior evolved for all four evolutionary conditions (Fig. 7). The experiments with a cooperative task led to a change in the relative performance of the four evolutionary conditions. Performance of homogeneous teams evolved with individual selection and homogeneous teams evolved with team selection was significantly higher than that of heterogeneous teams evolved with individual selection and heterogeneous teams evolved with team selection (all four $P < 0.001$), with the best fitness values in homogeneous teams up to 70% higher than those in heterogeneous teams. One possible reason is disparities in genome evaluation between homogeneous and heterogeneous teams (see next section). Another possible reason is that selection of good genomes could only happen indirectly in this task, which may have led to inefficient selection just as in heterogeneous teams evolved with team selection in task 1 (Section IV-B). This could be because fitness in this task was assigned to all team members, irrespective of their participation in the token foraging.

Performance of homogeneous teams evolved with individual selection and homogeneous teams evolved with team selection did not differ significantly ($P = 0.839$). This was because fitness in this task was assigned to all team members, irrespective of their participation in the token foraging. For the same reason, performance of heterogeneous teams evolved with individual selection and heterogeneous teams evolved with team selection did not differ significantly ($P = 0.365$).

D. Task 2—Cooperative Foraging: Disparities in Genome Evaluation and Credit Assignment

To test the hypothesis that the differences in performance of heterogeneous teams that evolved with individual selection and homogeneous teams that evolved with individual selection and team selection were caused by disparities in genome evaluation or by the fitness assignment to all team members, irrespective of their participation in the token foraging. For the same reason, performance of heterogeneous teams evolved with individual selection and heterogeneous teams evolved with team selection did not differ significantly ($P = 0.365$).

An electronic supplement for this paper is available online at http://lis.epfl.ch/documentation.php
Heterogeneous teams evolved with team selection performed worse than all other evolutionary conditions due to inefficient selection (all three $P < 0.001$).

E. Task 3—Altruistic Cooperative Foraging

Successful foraging behaviors evolved for all four evolutionary conditions (Fig. 9). Team performance in the altruistic cooperative foraging task was systematically lower than in the cooperative foraging task. This may seem surprising because the larger number of tokens in the arena increased the total number of fitness points available. A possible reason is that the increased number of tokens led to more clutter in the arena, which made successful token transport more difficult (see video supplied with supplementary material). Homogeneous teams achieved significantly higher fitness values than heterogeneous teams (all four $P < 0.001$). Possible reasons are disparities in genome evaluation and inefficient selection for the foraging of large tokens because fitness points gained from large tokens were assigned to all team members, irrespective of their participation in the token foraging (see next section).

Performance of homogeneous teams evolved with individual selection and homogeneous teams evolved with team selection did not differ significantly ($P = 0.310$). This was because selection probabilities for a given genome were again equal for both homogeneous evolutionary conditions.

looseness Performance of heterogeneous teams evolved with individual selection and heterogeneous teams evolved with team selection did not differ significantly ($P = 0.490$). However, the four evolutionary conditions resulted in different foraging strategies in this task (Fig. 10): While homogeneous teams evolved with individual selection and homogeneous teams evolved with team selection as well as heterogeneous teams evolved with team selection collected a significantly higher proportion of large tokens than small tokens (all three $P < 0.001$), heterogeneous teams evolved with individual selection collected a significantly higher proportion of small tokens than large tokens ($P < 0.001$). In comparison to the other three evolutionary conditions, heterogeneous teams evolved with individual selection collected significantly high proportion of small tokens (all three $P < 0.001$), but significantly low proportion of large tokens of all four evolutionary conditions (all three $P < 0.003$).

F. Task 3—Altruistic Cooperative Foraging: Disparities in Genome Evaluation and Credit Assignment

To test the hypothesis that the differences in performance of heterogeneous teams evolved with individual selection were caused by disparities in genome evaluation or by the fitness assignment to all team members, we performed two sets of additional experiments similar to those described in Section IV-D. First, we again corrected for the disparities in genome evaluation. However, correcting for this factor alone

---

Fig. 8. Task 2—Cooperative foraging without disparities in genome evaluation and without credit assignment problems. (a) Homogeneous teams evolved with one evaluation per team (instead of 10) and (b) heterogeneous teams evolved with 100 agents per population (instead of 1000). The performance of heterogeneous teams evolved with individual selection was higher than the performance of heterogeneous teams evolved with team selection, but did not reach that of homogeneous teams in this task. For boxplot explanations, see Fig. 5.
**Team fitness**

**Hom.**, **Ind. Sel.**

**Het.**, **Ind. Sel.**

**Het.**, **Team Sel.**

**0**

<table>
<thead>
<tr>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
</tr>
</thead>
</table>

**Fig. 9.** Task 3—Altruistic Cooperative Foraging. Left: Evolution of the best team fitness averaged over the best teams in 20 independent evolutionary runs over 300 generations. Right: The best team at generation 300 of each of the 20 independent experiments per evolutionary condition and per task was evaluated 1000 times. Homogeneous teams performed significantly better than heterogeneous teams. For boxplot explanations, see Fig. 5.

**Proportion of tokens collected**

**Hom.**, **Ind. Sel.**

**Hom.**, **Team Sel.**

**Het.**, **Ind. Sel.**

**Het.**, **Team Sel.**

**Small**

**Large**

**0**

<table>
<thead>
<tr>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
</tr>
</thead>
</table>

**Fig. 10.** Task 3—Altruistic cooperative foraging. The plot shows the average proportion of the six small tokens and four large tokens collected by the best teams at generation 300 for each of the 20 independent experiments and for each of the four evolutionary conditions. Heterogeneous teams evolved with individual selection pursued a different foraging strategy than teams of the other three evolutionary conditions, collecting very few large tokens but most small tokens.

**Fig. 11.** Task 3—Altruistic cooperative foraging without disparities in genome evaluation and without credit assignment problems. (a) Homogeneous teams evolved with one evaluation per team (instead of 10) and (b) heterogeneous teams evolved with 100 agents per population (instead of 1000). The performance of heterogeneous teams evolved with individual selection was higher than the performance of heterogeneous teams evolved with team selection, but did not reach that of homogeneous teams. For boxplot explanations, see Fig. 5.

In this second additional set of experiments, the performance of homogeneous teams that evolved with individual selection and homogeneous teams that evolved with team selection did not differ ($P = 0.133$). Heterogeneous teams evolved with team selection performed worse than all other evolutionary conditions due to inefficient selection (all three $P < 0.001$).

Importantly, the altruistic cooperative foraging task led to the evolution of a different foraging strategy in heterogeneous teams evolved with individual selection than in the other three evolutionary conditions (Fig. 10). A possible reason is that cooperation to collect large tokens now implied a cost for individuals. To test this hypothesis, we performed additional experiments with this evolutionary condition. First we repeated the experiments with a setup identical to that of task 3, i.e.,...
teams that evolved with team selection was not affected by the costs implied in large token foraging (see Figs. S5 and S6 in the supplementary material\(^5\)). This was because relative fitness differences between team members could not have an influence on the selection of genomes when individuals were genetically identical.

Foraging strategy in heterogeneous teams that evolved with team selection was not affected by the costs implied in large token foraging (see Figs. S5 and S6 in the supplementary material\(^5\)). This was because relative fitness differences between team members did not have an influence on selection of genomes when selection acted at the level of the team.

V. CONCLUSION

This paper provides an experimental demonstration of how the choice of genetic team composition and level of selection influences the performance of MAS in tasks with varying levels of cooperation that do not provide a benefit for specialization. We have identified three different types of multiagent tasks depending on the amount of cooperation required between team members. Our results demonstrate that different combinations of genetic team composition and level of selection lead to significant performance differences. No combination achieved optimal performance in all three types of task.

We have identified and studied three different types of multiagent tasks depending on the amount of cooperation required between team members.

In tasks that did not require cooperation, heterogeneous teams that evolved with individual level selection achieved the highest team performance. Team heterogeneity allowed evaluation of a high number of different genomes in parallel, and individual selection allowed efficient selection of good genomes. However, these teams performed poorly in tasks that required cooperation and in tasks with credit assignment problems.

For multiagent tasks that required cooperation, the highest team performance was achieved by homogeneous teams. These teams led to efficient cooperation between team members and they were not affected by credit assignment problems or costs associated with cooperation. Our results suggest that homogeneous teams are a safe choice in tasks that do not benefit from specialization when the requirements for agent cooperation are difficult to estimate. Compared to heterogeneous teams, homogeneous teams evaluate less genomes, which may result in premature convergence to suboptimal solutions. Our experimental results indicate that a simple way to prevent this problem is to use populations made of a large number of homogeneous teams.

Heterogeneous teams evolved with team selection were inefficient at selecting for good performance in all three types of tasks studied here and therefore cannot be recommended for cooperative tasks that do not require specialization.

Converging evidence towards these guidelines was recently found in a study on the evolutionary conditions for the emergence of communication in societies of robots with different

\(^5\)An electronic supplement for this paper is available online at http://lis.epfl.ch/documentation.php
morphologies and sensing abilities than those described in this paper [17].

However, it should be noted that evidence from both studies has two notable limitations. First, they did not address tasks that benefit from specialization in addition to cooperation. There is evidence that behavioral heterogeneity can lead to significant performance advantages for such tasks [29], [63], [85]–[88]. Second, they did not consider teams with intermediate genetic similarity. Biological research has shown that such teams can overcome individual fitness costs of cooperation [55], [89], thus combining the best of both worlds—enhanced genetic diversity with readiness to adopt altruistic behaviors. A good understanding of those conditions will require significant further research.

REFERENCES


Markus Waibel (M’08) received the Masters degree in physics from the Technical University of Vienna, Vienna, Austria, in 2003, and the Ph.D. degree from the Laboratory of Intelligent Systems (LIS), School of Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, in 2007.

He is currently with the LIS, EPFL. His research interest is in the evolution of cooperation in both natural and artificial systems. He has coordinated the popular “Talking Robots” and “Robots” podcasts and is a regular contributor to the IEEE robotics blog *Automaton*.
Laurent Keller received the B.Sc. degree in Biology, M.Sc. degree in Biology and Ph.D. degree in Zoology from University of Lausanne, Lausanne, Switzerland, in 1983, 1985, and 1989, respectively. He is a Professor of Ecology and Evolution, and Head of the Department of Ecology and Evolution, Biophore, University of Lausanne, Lausanne, Switzerland. His research interests include the principles governing the evolution of animal societies and the ecological and evolutionary consequences of social life. In addition to publishing several research papers, he has edited the books *Queen Number and Sociality in Insects* in 1993 and *Levels of Selection in Evolution* in 1999.

Dr. Keller was awarded the E. O. Wilson Naturalist Award in 2005.

Dario Floreano (SM’05) received an M.A. in visual psychophysics at the University of Trieste in 1988, an M.Sc. in Neural Computation from the University of Stirling in 1992, and a Ph.D in Cognitive Systems and Robotics from University of Trieste in 1995. He is an Associate Professor and the Director of the Laboratory of Intelligent Systems, School of Engineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. His research interest includes the analysis and synthesis of bio-inspired adaptive systems, such as biomimetic robotics, neuromorphic engineering, and artificial evolution. He has published two books on these topics: *Evolutionary Robotics* with Stefano Nolfi (Cambridge, MA: MIT Press, 2000) and *Bio-Inspired Artificial Intelligence* with Claudio Mattiussi (Cambridge, MA: MIT Press, 2008).