Diversity and Specialization in Collaborative Swarm Systems

Ling Li[1], **Alcherio Martinoli**[2], **Yaser S. Abu-Mostafa**[1]

1. California Institute of Technology, Pasadena, CA 91125, USA.
   Corresponding author: `ling@caltech.edu`
2. Swiss Federal Institute of Technology, CH-1015 Lausanne, Switzerland.

Abstract

This paper addresses qualitative and quantitative diversity and specialization issues in the framework of self-organizing, distributed, artificial systems. Both diversity and specialization are obtained via distributed learning from initially homogeneous swarms. While measuring diversity essentially quantifies differences among the individuals, assessing the degree of specialization implies to correlate the swarm's heterogeneity with its overall performance. Starting from a stick-pulling experiment in collective robotics, a task that requires the collaboration of two robots, we abstract and generalize in simulation the task constraints to k robots collaborating sequentially or in parallel. We investigate quantitatively the influence of task constraints and type of reinforcement signals on diversity and specialization in these collaborative experiments. Results show that, though diversity is not explicitly rewarded in our learning algorithm and there is no explicit communication among agents, the swarm becomes specialized after learning. The degree of specialization is affected strongly by environmental conditions and task constraints, and reveals characteristics related to performance and learning in a more consistent and clearer way than diversity does.

Keywords: collaborative swarm systems, distributed learning, specialization, diversity.

1 Introduction

Artificial swarm systems based on swarm intelligence (SI) consist of relatively simple autonomous agents. They are truly distributed, self-organized, inherently scalable since there is no global control or communication mechanism, and exploit an adequate balance between explorative and exploitative behavior for robustly facing changes in environmental or task conditions (Bonabeau et al., 1999).

Swarm systems can be homogeneous or heterogeneous. A homogeneous system consists of physically identical entities with the same hardware and software capabilities. A heterogeneous system may differentiate at different levels: at the hardware level, at the (controller) software level, or simply because each entity has a unique identifier. In this paper, we use software agents emulating real robots that differentiate exclusively at the controller level, in particular endowed with different control parameters.

Homogeneous systems represent a special case of heterogeneous ones. Depending on environmental and task constraints, a homogeneous solution may not be that achieving the best results. Learning, as an automatic way to adjust control parameters or select rules without a priori assuming the degree of swarm heterogeneity, represents an effective tool to explore not only homogeneous solutions (Hayes et al., 2003) but also heterogeneous ones (Murciano et al., 1997; Li et al., 2002). In this paper, we are interested in distributed learning, i.e., adaptation through learning occurs exclusively in single robot's controllers (and not, for instance, in an external supervisor unit).

However, depending on agents' capabilities in perception and communication, it may be extremely difficult for a distributed learning algorithm to discover (near-)optimal solutions at the swarm level. In addition to the inherently large search space characterizing a heterogeneous swarm, the credit assignment problem due to partial perception of agents drastically increases the difficulty of distributed learning. Solutions proposed in the literature of multi-agent learning can be roughly classified according to the type of reinforcement signal adopted, either local or global. The *local reinforcement signal* (Matarić, 1998; Parker & Touzet, 2000) rewards a single agent based on the local assessment about its contribution to the swarm performance. Although this type of reinforcement signal is immediate and exploits the inherent parallelism of the swarm, it just represents a noisy estimation of the swarm performance. The more limited and local the communication and perception capabilities (e.g., in extreme cases no communication at all and very short-range sensors) are, the higher the amount of noise is in the local assessment due to partial perception. On the contrary, the *global reinforcement signal* (Murciano et al., 1997; Versino & Gambardella, 1997; Hayes et al., 2003), which is often equivalent to the swarm performance, is stabler and more meaningful. However, this usually implies a reliable way to measure the swarm performance (e.g., a supervisor or a fully connected, fast communication network among agents) and a more difficult interpretation of the reinforcement signal at the agent level, especially in heterogeneous systems.

In this paper, we let the distributed learning algorithm explore heterogeneous solutions, aiming to improve the swarm performance. We consider different task constraints and types of reinforcement signals, and quantitatively measure diversity and specialization of a team of non-communicating agents. We support the discussion first with a concrete collaboration experiment concerned with pulling sticks and then with its generalized versions where the collaboration is extended to k sequential or parallel operations—with the analog of pulling longer or heavier sticks. We show that specialization can arise in all versions of experiments as a function of task constraints and environmental conditions no matter which type of reinforcement signal is used. As long as the diversity in agents brings advantage to the swarm performance, learning can drive the system to be specialized.

2 Diversity and Specialization

Traditionally, swarm systems have been classified on a bipolar scale as either heterogeneous or homogeneous depending on whether any of the agents differ from the others. This view is limiting because it does not permit a quantitative comparison between heterogeneous systems. Quantitative metrics of swarm diversity and specialization can enable the investigation of issues such as the impact of diversity on swarm performance and the impact of task constraints on specialization.

The essential idea behind the diversity measure is to cluster similar agents according to a problem-specific difference measure and look at the pattern they form in the feature space. After some preliminary tests where we used a heuristic criterion to select the "optimal" clustering and adopted the number of clusters as the diversity measure (Li, 2002), we adopted Balch's *social entropy* (Balch, 1998) as the diversity measure for the stick-pulling experiments. Based on Shannon's information entropy, Balch's social entropy makes a meaningful and stable measure by incorporating details about the feature space such as the spatial distribution of the clusters.

Specialization means more than just being diverse. While diversity means difference among individuals no matter whether the difference is good or bad in respect to the swarm performance, specialization, with the definition "structural adaptation of a part to a particular function," also means adaptation in order to fit. When diversity is obtained via an iterative process such as learning or evolution, other reasons (e.g., noise in the replication mechanism) can also cause the system to become diverse. However, a system becomes specialized when, given specific constraints of viability or survival at the agent level, its diversity is caused for better performance. Accordingly, a specialization metric should measure the part of diversity that enhances the performance.

When looking at a swarm system statically, it is impossible to identify the part of diversity that corresponds to the performance improvement. We have to put the system into a dynamic process where its performance and diversity can change and interact. If the performance generally increases with higher diversity, the system benefits from being more diverse than the initial status, and the degree of specialization should increase accordingly; otherwise, if the greater diversity does not help the performance, the degree of specialization should decrease. That is, specialization can be measured along a dynamic process as a result of the correlation between the diversity and the performance. If we assume the system starts from a homogeneous setting with no diversity or specialization, and the diversity d and the swarm performance r changes with time as correlated random variables, the correlation coefficient between d and r acts naturally as the percentage of specialization in diversity. To put this in a formula, the degree of specialization can be defined as

$$s = \text{corrcoef}(d; r) \times d. \tag{1}$$

Note that our specialization measure s is negative when d and r are negatively correlated.

3 Stick-Pulling Experiments

Ijspeert et al. (2001) investigated collaboration in teams of non-communicating robots engaged in a stick-pulling experiment (Figure 1, left). We call their experiment the original one since we will abstract and generalize it later in Subsection 3.2.

3.1 Original Stick-Pulling Experiment

In the original experiment, robots equipped with gripper turrets and proximity sensors search a circular arena and pull sticks out of the ground. The stick length has been chosen so that a single robot is incapable of pulling a stick out completely on its own, but collaboration between two robots is sufficient for this task. Each robot is characterized by a *gripping time parameter* (GTP) which is the maximal length of time that a robot waits for the help of another robot while holding a stick.

The behavior of a robot is determined by a simple program (Figure 1, right). The default behavior is searching for sticks, i.e., wandering in the arena until an object is detected. If a stick is detected, the robot pulls it up and determines whether another robot is already holding it by measuring the elevation speed of the gripper arm. If the elevation is fast, there is no other robot holding the stick and we call such a grip *grip1*. Otherwise, the robot assumes that another robot is already holding the stick and therefore “braking” the elevation. Such a grip is named *grip2*.

After a robot makes a grip1, two cases can occur: either a second robot helps the first one before the GTP expires (we call this a *successful collaboration*) or the first robot times out and resumes the search for sticks. The specific values of GTPs play a crucial role in the overall *stick-pulling rate* (defined as the number

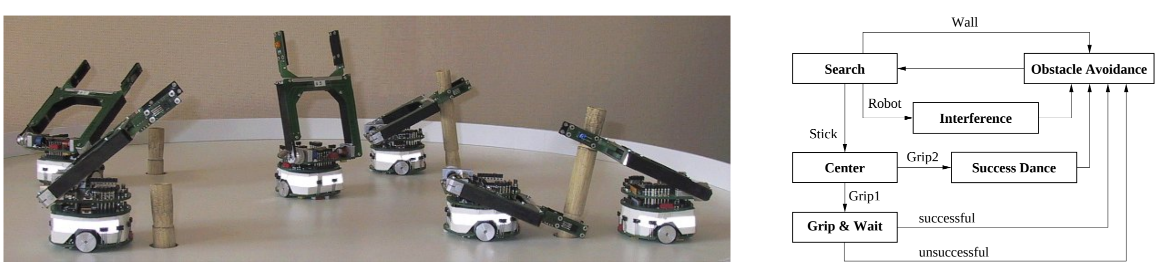

Figure 1: Left: Physical set-up for the stick-pulling experiment. Right: FSM representing the robot's controller. Transitions between states are triggered by sensory measurements.

of sticks pulled out per unit time) which is the metric adopted in all previous papers[5] (Ijspeert et al., 2001; Lerman et al., 2001; Li et al., 2002) and this paper for the swarm performance. To ensure the stick-pulling rate is reliably measured, experiments usually take a long time and a stick will be inserted back by the experimenter after it is completely pulled out.

We use the microscopic model developed in (Ijspeert et al., 2001) as the simulation platform, which represents agents as separate probabilistic finite-state machines (PFSM). The flowchart of a PFSM is based on the blueprint of the corresponding real robot controller and its transition probabilities are computed using simple geometric considerations and systematic experiments with one or two real robots. Unlike macroscopic models (see for instance (Lerman et al., 2001; Martinoli & Easton, 2002) for the same experiment) which intrinsically assume agents can be clustered into certain castes, microscopic models allow us to study issues related to distributed learning and specialization since each agent is a separate PFSM. Furthermore, in contrast to other agent-based models, the way this model is constructed allows for quantitatively accurate predictions while being four or five orders of magnitude faster than other popular simulation tools such as sensor-based embodied simulations (Ijspeert et al., 2001). Therefore, although we have not tested our results using real robots or realistic simulations, we believe that their validity is not limited to abstract agents.

3.2 Generalized Stick-Pulling Experiments

The strict collaboration property of the stick-pulling task has a major influence on swarm diversity and specialization. In order to emphasize this effect, we abstract and generalize the original experiment so that a successful collaboration requires now k (> 2) robots instead of just two.

Sequential Collaboration: Pulling Longer Sticks One way to extend the original experiment is to assume longer sticks so that one robot can only pull a stick up by $1/k$ of its length. k consecutive grips, which may be called grip1, grip2, $\ldots$, and grip$k$, respectively, are thus needed for pulling out a stick entirely. If the robot currently holding the stick times out, it will drop the stick so that further robots will have to start over from grip1. We call this type of collaboration required for pulling longer sticks *sequential collaboration*. Note that we do not really need more than two robots in order to complete the task. Theoretically, two robots with very large GTPs are able to pull out sticks of any length but inefficiently, if they help each other alternately.

Parallel Collaboration: Pulling Heavier Sticks Another way to extend the original experiment is to suppose the sticks are shorter but heavier so that one robot is too weak to lift a stick up. Exact k robots are needed simultaneously to lift a stick and pull it out. When a robot finds a stick, it grips the stick until timing out or until there are enough robots to lift the stick, whichever comes earlier. Robots do not reset their timers when a new robot joins the pulling. Distinguished from the sequential case, unless all the robots currently holding the same stick time out, the pulling process need not to be restarted from scratch. We call this type of collaboration *parallel collaboration*.

3.3 Learning Algorithm

We proposed and tested in (Li et al., 2002) an adaptive line-search algorithm and found that the algorithm could achieve near-optimal performance in the original stick-pulling experiment under different conditions. In contrast to a gradient descent method, this algorithm neither requires the derivative to be calculated nor assumes continuity in the search space. In this paper, we use the same algorithm for both the original and the generalized stick-pulling experiments.

[5]The *collaboration rate* (the number of successful collaborations per unit time) was in fact used in the previous papers. It is equivalent to the stick-pulling rate when exact one successful collaboration is required for a stick pull-out.

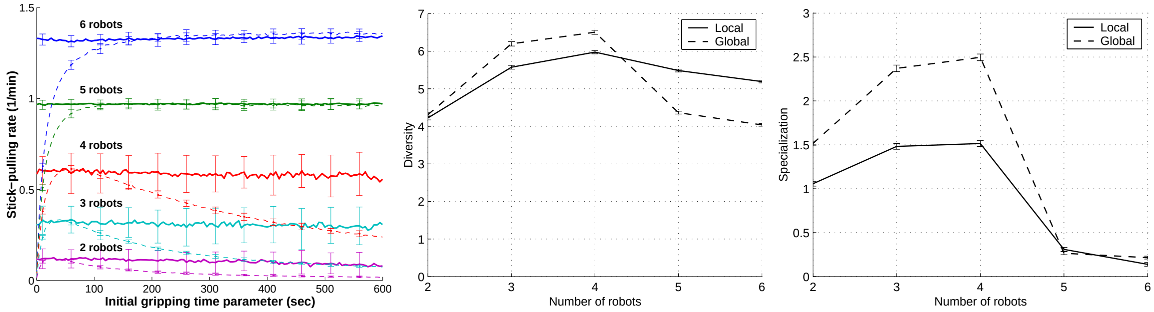

Figure 2: Results of the original stick-pulling experiment. Left: The dashed curves represent the performance of homogeneous teams with a fixed GTP; the solid curves show that of heterogeneous teams after learning under local reinforcement signal. Middle: The diversity under different reinforcement signals. Right: The specialization under different reinforcement signals.

We use both types of reinforcement signals with the learning algorithm. The local reinforcement signal rewards an agent when it makes a successful collaboration, i.e., when it completely pulls out a stick or passes the stick to another robot. The global reinforcement signal is the swarm performance. The two types of reinforcement signals “align” well in the original experiment as well as its parallel extension since a successful collaboration means exactly a stick pull-out and vice versa. However, in sequential cases, a successful collaboration only contributes to but may not finally result in a stick pull-out, and without a supervisor or explicit communication, a robot will never know its true contribution unless it does the final grip. Thus the local reinforcement signal in sequential cases is not aligned with the global one.

4 Results

All the experiments we conducted started from a homogeneous system, i.e., a same initial GTP for all agents. During the experiments, agents could iteratively adapt their GTPs using either the local or the global reinforcement signal. The experiments lasted long enough for the learning to stabilize. Swarm performance and diversity were recorded along the experiments using a time window so that specialization could be measured via formula (1). We simulated 50 runs for each initial GTP and plotted the mean diversity and specialization over the runs. The error bars in all diversity and specialization figures represent one standard deviation from the mean values.

In (Li, 2002), we suggested to use a difference measure of logarithmic form since both the performance and the logarithm are less sensitive to GTP changes when GTP is large. That is, for two agents with GTPs g_1 and g_2 respectively, the difference between them is $|\log g_1 - \log g_2|$. This difference measure is used in all of our experiments.

4.1 Results of the Original Stick-Pulling Experiment

We started with the original stick-pulling experiment using the same settings as in (Ijspeert et al., 2001; Li et al., 2002), i.e., 2 to 6 robots and 4 sticks in an arena of 40 cm in radius. The learned performance under the local reinforcement signal contrasted with the performance of a homogeneous team without learning is shown in the plot on the left of Figure 2.

The homogeneous team with a fixed GTP exhibited quite different behaviors depending on the robot/stick ratio. When there were more robots than sticks, the stick-pulling rate increased monotonically with the GTP until reaching a plateau corresponding to the optimal rate for homogeneous teams. In other words, since there were always robots "free" to help, waiting very long was a good strategy for robots holding sticks. On the other side, when the number of robots was equal to or smaller than that of sticks, waiting in vain for a very long time may generate *deadlock* situations where every robot holds a different stick and waits for help. Previous research showed that specialization was desired particularly in this situation (Ijspeert et al., 2001; Li et al., 2002).

The stick-pulling rate of the learned system instead consistently achieved the same level independent of the initial GTP and almost always outperformed the rate obtained by the homogeneous team without learning. We also tested learning with the global reinforcement signal. Probably due to high alignment between the local and the global signals under the current task constraints, we did not observe significant difference in the learned performance under these two types of signals.

The plot on the right of Figure 2 shows that specialization became much smaller for 5 and 6 robots than for 2–4 robots. This validates the deadlock phenomenon we just discussed, i.e., diversity is good for the performance when there are equal or less number of robots than sticks, and becomes less relevant with the performance when there are more robots than sticks. The diversity measure (Figure 2, middle) gave flatter curves and by itself cannot show this phenomenon clearly.

Since the local reinforcement signal is noisier than the global one, we expect that under the global reinforcement signal truly specialized robots generate a larger portion of the diversity. This is validated in Figure 2 since the diversity under the global reinforcement signal dropped faster than that under the local reinforcement signal when the specialization was less relevant.

4.2 Results of the Generalized Stick-Pulling Experiments

In order to accommodate more robots required by the generalized experiments, we used a large arena of 80 cm in radius, 16 sticks, and 6 to 24 robots. We simulated the generalized experiments with k from 3 to 5. Probably due to the same "alignment" reason in the original experiment, no significant difference in performance was spotted between the local and the global reinforcement signals in parallel cases. However, in sequential cases, the local reinforcement signal gained a small performance advantage over the global one for almost all experimental settings.

Before looking at the specialization results (Figure 3), we had expected that the specialization in parallel cases would be higher than that in sequential cases.[6] However, that happened only with large number of robots (say, 18) and the global reinforcement signal. An investigation of the learned GTPs shows that when the number of robots is small in parallel cases, all robots have similar GTPs ($\sim 300\,\text{s}$) and the diversity is low. This gives us hints about the seemingly weird phenomenon.

We define the *deadlock threshold* as the maximal number of robots that could still incur deadlock. When there are t sticks in the arena, the threshold is t in sequential cases and $(k-1)t$ in parallel cases. Our experience with the original experiment made us believe that specialization is high any time the number of robots is less than the deadlock threshold, which is not always true. Just as in a company having much more jobs than employees, when the deadlock threshold is much higher than the number of robots, each robot tends to have multiple roles, as every employee has to take multiple jobs. Since a robot has only one GTP value, trying to specialize into too many directions just makes all GTPs similar and results in a low diversity, especially when k is large in parallel cases.

With the global reinforcement signal, when the number of robots is larger than the deadlock threshold, the decreasing of specialization was again observed.[7] What was initially unexpected is that specialization achieved its maximum when the number of robots was measurably lower than the threshold. However, seeing

[6] Our arguments were: (a) In sequential cases, the requirement for robots doing grips before grip$k$ is similar—their GTPs are large enough for the next robot to come and take over. In parallel cases, k different GTP values may be instead required—robots doing grip1 need the largest GTP and robots doing grip$k$ need the smallest GTP. (b) With the same number of sticks, the parallel collaboration essentially requires more robots working simultaneously. We know from the original experiment that specialists may arise if there are insufficient robots compared with sticks.

[7] For parallel cases, since the threshold is much higher, we verified this with 2 sticks, 4 to 9 robots, and $k = 4$.

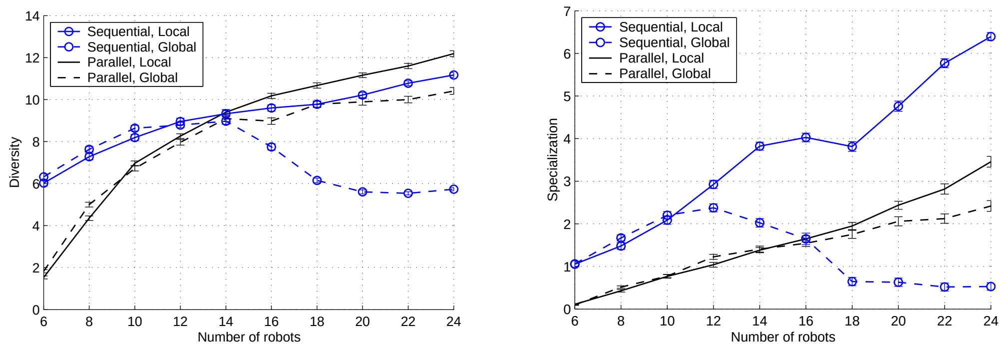

Figure 3: Diversity and specialization in the generalized experiments with $k = 4$.

that the deadlock threshold is a pessimistic estimation since the agents cannot have infinitely large GTPs, the "real" threshold should be smaller.

5 Conclusions

This paper presented our initial effort to measure specialization in collaborative swarm systems. Specialization is a mixed concept of both diversity and adaptation. We define specialization as the part of diversity that is induced by the need of performance improvement. Our experiments with the original and generalized stick-pulling experiments showed that specialization was more consistent and meaningful than diversity when properties related to performance and learning were under study. Our results validated some of our intuitions about specialization in these collaborative experiments but also revealed some properties that we at first did not see.

Our specialization measure depends heavily on the underlying dynamic process. Different learning algorithms might result in different specialization values even when the final learned systems are the same. Future work will be making the specialization measure more independent of the choice of the learning algorithm, or more generally speaking, of the dynamic process in which diversity and swarm performance interact.

Acknowledgements

This work has been mainly supported by the Caltech Center for Neuromorphic Systems Engineering under the US NSF Cooperative Agreement EEC-9402726 and the Northrop Grumman Corporation Foundation. Alcherio Martinoli is currently sponsored by a Swiss NSF professorship.

References

Balch, T. 1998. *Behavioral Diversity in Learning Robot Teams.* Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA.

Bonabeau, E., Dorigo, M. & Theraulaz, G. 1999. *Swarm Intelligence: From Natural to Artificial Systems.* Oxford University Press, New York.

Hayes, A. T., Martinoli, A. & Goodman, R. M. 2003. Swarm robotic odor localization: Off-line optimization and validation with real robots. *Robotica* **21**(4): 427–441.

Ijspeert, A. J., Martinoli, A., Billard, A. & Gambardella, L. M. 2001. Collaboration through the exploitation of local interactions in autonomous collective robotics: The stick pulling experiment. *Autonomous Robots* **11**(2): 149–171.

Lerman, K., Galstyan, A., Martinoli, A. & Ijspeert, A. J. 2001. A macroscopic analytical model of collaboration in distributed robotic systems. *Artificial Life* **7**(4): 375–393.

Li, L. 2002. *Distributed learning in swarm systems: A case study.* Master's thesis, California Institute of Technology, Pasadena, CA.

Li, L., Martinoli, A. & Abu-Mostafa, Y. S. 2002. Emergent specialization in swarm systems. Pages 261–266 in: *Intelligent Data Engineering and Automated Learning — IDEAL 2002* (H. Yin et al., eds), vol. 2412 of *Lecture Notes in Computer Science*, Springer-Verlag.

Martinoli, A. & Easton, K. 2002. Modeling swarm robotic systems. Pages 297–306 in: *Experimental Robotics VIII* (B. Siciliano & P. Dario, eds), vol. 5 of *Springer Tracts in Advanced Robotics*, Springer-Verlag.

Matarić. M. J. 1998. Using communication to reduce locality in distributed multi-agent learning. *Journal of Experimental and Theoretical Artificial Intelligence* **10**(3): 357–369.

Murciano, A., Millán, J. del R. & Zamora, J. 1997. Specialization in multi-agent systems through learning. *Biological Cybernetics* **76**(5): 375–382.

Parker, L. E. & Touzet, C. 2000. Multi-robot learning in a cooperative observation task. Pages 391–401 in: *Distributed Autonomous Robotic Systems 4* (L. E. Parker et al., eds), Springer-Verlag.

Versino, C. & Gambardella, L. M. 1997. Learning real team solutions. Pages 40–61 in: *Distributed Artificial Intelligence Meets Machine Learning: Learning in Multi-Agent Environments* (G. Weiß., ed.), vol. 1221 of *Lecture Notes in Artificial Intelligence*, Springer-Verlag.