Research Statement

Paulina Varshavskaya

Imagine a team of autonomous robots clearing out a disaster scene, perhaps that of a collapsed building. Suppose some of them have figured out how to coordinate their actions in order to lift a large, heavy concrete slab of a particular shape. What does that tell nearby robots who are attempting to lift another slab of a different shape, size and weight? Observe a "wisdom of crowds" phenomenon in a persistent group of individuals: a shoal of fish moving cohesively to avoid a predator or to find a food source, or an economy of human players solving the problem of continuous resource and goods allocation. As an alert and learning participant in such a group, how can I use my observations of what others are doing to inform my own choices and actions?

These questions exemplify my main research interest, which is studying and developing distributed intelligent systems that adapt and learn. Our world teems with such systems: be they natural as locust swarms, societal as organizations and markets, artificial as computer and traffic networks or robot teams, or hybrid. Much about the learning capabilities of these systems is still unknown, yet we can think of their essence as continual decision-making by individuals under uncertainty with limited, heterogeneous information while trading off a number of goals. How does information flow through these groups and affect the collective outcome? How should individuals make decisions optimally in these groups? How do group properties and phenomena affect the optimality of decisions? How is group learning or optimality even to be defined and measured? What relationship should individual goals have to the group outcome? Drawing on elements of Markov decision processes (MDPs) and reinforcement learning, game theory and mechanism design, and mathematical models of distributed systems in biology, I want to develop a research program which addresses the broad questions of learning in decentralized, multi-agent systems. How does a group of agents with limited individual resources and intelligence collectively acquire new knowledge and behavior? How can that process be modeled in natural systems? How can it be directed and optimized in engineered ones?

I have begun to address these questions through the design and development of distributed robot systems and distributed machine learning algorithms to control such systems. An early example was a year-long collaboration with Joerg Conradt from ETH Zurich, where I designed and built a novel modular serpentine robot, the first such robot to be controlled in a fully distributed way by coupled central pattern generators in each module [1]. Our work showcased the robustness of coupled distributed control, which was inspired by the neural control of lamprey locomotion. For my PhD thesis work, I developed distributed reinforcement learning algorithms to automate the generation of controllers for self-reconfiguring modular robots (SRMRs). I designed and implemented a class of algorithms for learning by policy search in this formally difficult, partially observable domain; they succeded in learning locomotion gaits where standard reinforcement learning algorithms, which make the full observability assumption, failed. The issues of local optima and scalability were addressed both by introducing local communication schemes, and by leveraging group properties through incremental learning with progressively larger numbers of modules [2]. A particularly successful approach that I developed combined local agreement and distributed policy search for sharing both local rewards and local experience among modules, which resulted in a tenfold increase of learning speed [3].

Intelligent control algorithms need to be designed for and validated in embodied systems in the physical world; and I intend to pursue this validation in future research. My extensive background in robotics makes me well-prepared to engage in this research. For example, I co-designed the original prototype AMOUR [4] — an underwater robot which Iuliu Vasilescu and I had conceived as a potentially self-replicating machine, and which has since become a mobile part of an underwater sensor network. I was part of the large team integration effort on the Cardea mobile manipulation project [5], where I contributed an early system for visual processing, and parts of the overall behavior-based control, as well as some hardware development.

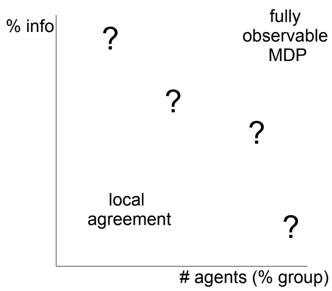
The distributed agreement-while-learning algorithms from my thesis work provide a starting point in answering a broad and exciting two-fold research question: how can learning individuals with limited resources benefit from being part of a group? And how can group goals be achieved by distributing computation and learning among individuals? Agreement algorithms give a simple procedure for locally averaging experience among cooperating agents which would ideally be executing the same policy. But in many distributed systems, both natural and artificial, individuals may be required to behave differently for optimal results. I propose the following research program to extend these ideas towards diversity of information and behavior.

Social reinforcement learning Social learning is the ability of an individual to derive some understanding of the world or the effects of its own behavior through observing other individuals, called *model* agents. In the general setting there should be minimal restrictions on model agents: they are not required to demonstrate any specific behavior, nor to communicate with the learner; they may be competing with the learner, or trying to coordinate their behavior with the learner or others. While learning by imitation of a specific teacher has been addressed in the literature, a theory of learning under the minimal constraints above is still missing. I want to develop a general theory of optimal decision-making when social information is available to individuals and rewards are delayed. In current research, I am addressing the question of the utility of social information, and specifically the utility of agents as models for each other, for instance by defining agent-value functions by analogy with state and action-value functions commonly used in reinforcement learning. Taking this notion further will result in a social reinforcement learning framework with a new layer of decision-making relating to the possibility of imitation in addition to exploration and exploitation.

Algorithms and control of robot groups In many robotics applications, from space exploration to search and rescue, a team of robots will do better than a single robot. When a problem is distributed among a team, we can parallelize search and coverage, increase system robustness to random failures, and benefit from the potential for adaptation and enhanced power that is inherent in the group. In the search and rescue scenario, where a single robot may not be able to both lift a heavy slab in one place, and burrow into rubble to look for victims in another, a team could coordinate effort to perform both tasks, perhaps simultaneously. However, distributed algorithms for group coordination are notoriously difficult to design. I am therefore interested in developing distributed learning algorithms for control of robot groups. Theoretical advances in social learning will enable us to derive such adaptive distributed algorithms, which will be validated both in physically embodied and in simulated robotic systems on benchmark problems in multi-agent reinforcement learning. Social facilitation, when harnessed according to utility theory, should enhance and speed up the learning process and give better performance than existing distributed learning algorithms demonstrate. I am especially interested in applying these algorithms to the development of embodied systems to solve difficult engineering problems: autonomous construction, sensor networks, self-organizing machines.

Modeling of natural groups Social learning phenomena occur in many types of persistent groups in nature and society. A general theory of social learning will enable us to better predict behavior in such groups, and may even inform policy. I am therefore interested in using theoretical advances in understandng social learning for modeling natural group behavior. In a current collaboration with biologists, we are looking at the effects of social learning on motion patterns of fish shoals. I am also developing new collaborations between computer scientists and economists to tackle social information flow in groups where learning by imitation can be detrimental to group welfare (recent global financial disaster has been partly blamed on mistaken local social imitation). I am interested in pursuing these and other cross-disciplinary projects to arrive at a better understanding of group learning phenomena in all their diversity. Existing approaches either model only simple within-group interactions and assume no individual learning capacity, or focus on sophisticated individual learning to the exclusion of group goals. I believe that combining both approaches will lead to better models and more accurate predictions of both individual and collective behavior in natural groups.

Distributing group learning goals Another big question in distributed learning is: how does the group-level goal relate to individual goals? How can we distribute computation across individuals in an intelligent, optimal way? Consider a space of problem formulations which describes, loosely speaking, "how much I know about how many" as a learning participant in a group. We have seen systems where there is no information exchange (at the origin of this space), and those modeled as MDPs making the very unrealistic full observability assumption. In my thesis work, I have addressed a small region of this space with local agreement algorithms combined with policy search in partially observable problems such as SRMR locomotion: individual modules have access to a local estimate of their neighborhood's reward, or reward and history statistics. However, there has to date been no systematic exploration of this information space in terms of the kinds of algorithms and relationships between group and individual goals that it is likely to generate. I would like to engage in such an exploration.



There is great potential for significant and broad impact of this research program. In general, heterogeneity of information and knowledge across individuals is pervasive in the world. When should you figure things out on your own and when should you ask for a help? Who do you ask? When, if ever, do you volunteer to help somebody else? How does information travel in networks of agents trying to optimize by balancing exploration, exploitation and imitation? Answering these questions will lead to a greater understanding of existing natural and societal group learning phenomena, from persistent groups of moving animals, to the information and decision-making waves traveling through markets and economies. The results will also be useful for applications in robotics (distributed manipulation, decentralized control of many degrees of freedom, sensor networks, and swarm robotics), computer network systems, mobile sensor networks, or teams of cooperating heterogeneous agents.

References

[1] J. Conradt and P. Varshavskaya. Distrubuted central pattern generator control for a serpentine robot. In *13th Int. Conf. on Artificial Neural Networks (ICANN) Supplement. Proc.*, Istanbul, Turkey, June 2003.

[2] P. Varshavskaya, L.P. Kaelbling, and D. Rus. Automated design of adaptive controllers for modular robots using reinforcement learning. *International Journal of Robotics Research*, 27(3-4), Mar-Apr 2008.

[3] P. Varshavskaya, L.P. Kaelbling, and D. Rus. Efficient distributed reinforcement learning through agreement. In *Proc. 9th Int. Symposium on Distributed Autonomous Robotic Systems (DARS)*, Tsukuba, Japan, Nov 2008.

[4] I. Vasilescu, P. Varshavskaya, K. Kotay, and D. Rus. Autonomous modular optical underwater robot (AMOUR): Design, prototype and feasibility study. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA)*, Barcelona, Spain, Apr 2005.

[5] R.A. Brooks, L. Aryananda, A. Edsinger, P. Fitzpatrick, C.C. Kemp, U.-M. O'Reilly, E. Torres-Jara, P. Varshavskaya, and J. Weber. Sensing and manipulating built-for-human environments. *International Journal of Humanoid Robotics*, 1(1), Mar 2004.