Optimal Cooperation, Communication, and Learning in Decentralized Systems

Aditya Mahajan (McGill University),
Maxim Raginsky (University of Illinois at Urbana-Champaign),
Demosthenis Teneketzis (University of MIchigan),
Serdar Yüksel (Queen's University)

October 12 – October 17, 2014

The problem of optimal decision-making in decentralized systems arises in different application domains, including smart grids, cyber-physical systems, communication networks, machine learning, and information processing in organizations. Traditionally, these application domains have been investigated by different research communities, and each community has developed its own set of mathematical tools and theories to address optimal decentralized decision making. This workshop was the first of its kind, where researchers from these different communities were given an opportunity to be exposed to new ideas, angles, and perspectives on decentralized decision-making.

1 Overview of the Field, Recent Developments and Open Problems

Cooperation and coordination in decentralized systems. In decentralized systems, no decision maker (DM) knows the information known to all other DMs, yet all DMs must cooperate to achieve a common, system-wide objective. Multiple approaches have been used in the literature to achieve cooperation and coordination; they include: (a) Identifying optimality conditions so that a solution can be obtained by all DMs based on their local information (either using mathematical programming or dynamic programming); (b) Identifying projected sub-problems that are solved at each DM and iterating after exchanging information (either through a pricing mechanism or through explicit data communication subject to constraints) to reconcile the results.

Role of communication in decentralized systems. Communication or information-exchange is an important aspect of decentralized decision making because of the following: (a) Communication generates common knowledge among DMs. Such a common knowledge is useful for dynamic programming (see the previous bullet) and learning (see the next bullet). (b) When the DMs have an incentive to communicate, the global optimization problem is usually non-convex; while if such an incentive is absent, e.g., in static and partially nested teams, the global optimization problem may be convex. Thus, on one hand, the incentive to communicate makes the decentralized optimization problem harder. On the other hand, the presence of communication sometimes facilitates a dynamic programming decomposition, e.g., in partial history sharing. (c) Economic and technological constraints often impose a restriction on information exchange, which in turn imposes restrictions on the solution approaches.

Learning in decentralized systems. When the DMs communicate, they exchange information in order to reduce their uncertainty. This process of uncertainty reduction is generically referred to as learning. Existing research on learning in decentralized systems is along two complementary directions: (a) Bayesian learning (and its variants), which models the process by which the DMs form and refine their probabilistic beliefs about other DMs (including their knowledge, their strategies, etc.) and about the overall state of nature relevant to the problem at hand, assuming that all DMs conform to certain axioms of rational behavior; and (b) Non-Bayesian learning or learning by boundedly rational agents, which model the process of learning in repeated or sequential situations in the presence of various resource and complexity constraints. Bayesian learning describes the situation in which the DMs know the system model, so all the uncertainty arises due to the presence of other DMs, as well as due to local nature of communication; by contrast, bounded rationality deals with situations in which the system model is not completely known, so the DMs need to learn the model based on their local observations.

2 Presentation Highlights

Each day of the workshop featured several 40-minute technical talks, as well as several 50-minute overview talks. Most talks on each day were centered around one broad theme.

2.1 Day 1: Learning, stability, and games in multiagent Markovian environments

Overview talk – VIKRAM KRISHNAMURTY: *Social learning and active sensing.* Vikram opened up the workshop with an overview talk about structural results for partially observed Markov decision processes in multi-agent systems when individual agents perform *social learning*, i.e., they aggregate their private noisy signals about the global system state with the signals they receive from their neighbors in the network. He illustrated the key ideas using three specific examples: (1) constrained optimal social learning problem, where the onset of herding (i.e., agents exhibiting preference for nearly the same action) is delayed by agents sharing full information; (2) change detection when individual agents perform social learning; and (3) some recent (but incomplete) work by Vikram and collaborators on social networks where diffusion approximations are made resulting in a controlled Markovian system.

TARA JAVIDI: *Noisy Bayesian active learning.* Following the theme of learning in a Markovian environment, Tara discussed the problem of noisy Bayesian active learning, where we are given a class of functions and a sample space. An unknown function in this class assigns a label to any sample in the sample space, and the result of a label query on any sample is corrupted by independent stochastic noise. The goal is to identify the function that generates the labels with high reliability using as few label queries as possible, by selecting the queries adaptively in a strategic manner. Previous work on Bayesian active learning has revolved around a specific algorithm, the so-called Generalized Binary Search (and its variants for the noisy case), and provided upper bounds on the number of queries required by these sampling strategies. Tara first presented a converse result, characterizing the smallest number of queries that must be made by *any* adaptive algorithm as a function of the structure of the function class, the noise model, and the required reliability. The proof of this result relied on a blend of techniques from statistics, stochastic control theory, and information theory. She then revisited Generalized Binary Search and similar schemes in light of the converse result and showed that, in general, these schemes are suboptimal. She presented and analyzed an alternative strategy for sample collection based on DeGroot's notion of information utility: at each time step, this strategy queries the label of a sample which maximizes the so-called Extrinsic Jensen–Shannon divergence, which was recently introduced by Tara and her collaborators.

Overview talk – TAMER BAŞAR: *Stochastic dynamic games and intricacy of information structures.* In spite of decades-long past research activity on stochastic dynamic/differential games, there still remain some outstanding fundamental questions on the existence, uniqueness, and characterization of non-cooperative equilibria when players have access to noisy state information. Tamer's overview talk identified these questions, along with the underlying challenges, and addressed a number of them within specific contexts. One of the questions is related to the notion of certainty equivalence (CE), which is well-understood in single-player games (that is, stochastic control problems) but not in stochastic dynamic games, even if all players receive state information through a common (noisy) channel. Following a general overview of CE in stochastic control, Tamer discussed it in the context of two-player zero-sum stochastic differential/dynamic games (ZSSDGs) when the players have access to state information through a common noisy measurement channel — in both continuous time and discrete time. For the discrete-time case, the channel is also allowed to fail sporadically according to an independent Bernoulli process, leading to intermittent loss of measurements, and the players are allowed to observe past realizations of this process. Within this framework, Tamer presented a complete analysis of a parametrized two-stage stochastic dynamic game in terms of existence, uniqueness and characterization of saddle-point equilibria (SPE). Tamer then used the insight provided by the analysis of this game to obtain SPEs for three classes of differential/dynamic games: (i) linear-quadratic-Gaussian (LQG) zero-sum differential games with common noisy measurements, (ii) discrete-time LQG zero-sum dynamic games with common noisy measurements, and (iii) discrete-time LQG zero-sum dynamic games with intermittently missing perfect state measurements. In all of these cases, CE is a generalized notion, requiring two separate filters for the players, even though they have a common communication channel. There was a brief discussion of some extensions, such as (i) multi-player ZSSDGs, with teams playing against teams, where agents in each team do not have identical information, and (ii) nonzero-sum stochastic differential games, with asymmetric information among the players. The overview talk was concluded with a discussion of the mean-field approximation and the simplifications it brings at both the conceptual and analytical levels.

ALI BELABBAS: *Hurwitz graphs: algorithmic and probabilistic aspects.* Design of distributed control systems involves careful balancing of the complexity of the information structure against such desirable properties as stability, controllability, etc. Ali presented a mathematical formalization of this trade-off in the context of continuous-time distributed linear systems. The information structure (communication pattern between agents) in such a system is described by the locations of the nonzero entries in the system matrix. The question is: given the desired locations of the nonzero entries, does there exist at least one matrix that conforms to a desired information structure and is also *Hurwitz stable* (all of its eigenvalues have strictly negative real parts). This motivates the definition of a *Hurwitz space* as a vector space of square matrices that contains Hurwitz matrices. The talk focused on the design of Hurwitz vector spaces. After a brief overview of the main known results in the area, Ali presented new sufficient conditions guaranteeing that a vector space containing a Hurwitz subspace is itself Hurwitz. In the second part, he provided polynomial-time algorithms based on these sufficient conditions. He concluded the talk by discussing probabilistic aspects of Hurwitz spaces, where the locations of nonzero entries are chosen at random, and connected the emergence of Hurwitz property to monotone threshold phenomena in random graphs.

MICHEL DE LARA: *Smart power systems, renewable energies and markets: the optimization challenge.* Michel's talk was focused on concepts, models, and mathematical and numerical methods allowing to formalize and to treat issues concerning sustainability and precaution in the management of natural resources and of the environment. Specifically, he discussed mathematical concepts, such as controlled dynamical systems in discrete time, equilibrium and stability, viability and invariance, intertemporal optimality, stochastic and robust control; specific analysis and design methods, such as linearization, maximum principle, and dynamic programming; and simulation tools and algorithms. Throughout the talk, he focused on examples related to the exploitation of renewable or exhaustible resources and to greenhouse gases mitigation and proposed a common framework of sequential decision-making in this context.

SEAN MEYN: *Randomized policies for mean field control with application to automated demand response.* Sean's talk continued with the same theme as Michel de Lara's talk. Renewable energy sources such as wind and solar power have a high degree of unpredictability and time-variation, which makes balancing demand and supply challenging. One possible way to address this challenge is to harness the inherent flexibility in demand of many types of loads. At the grid-level, ancillary services may be seen as actuators in a large disturbance rejection problem. Sean argued that a randomized control architecture for an individual load can be designed to meet a number of objectives: The need to protect consumer privacy, the value of simple control of the aggregate at the grid level, and the need to avoid synchronization of loads that can lead to detrimental spikes in demand. He described new design techniques for randomized control of such systems, following a three-step program:

1. A parameterized family of average-reward MDP models is introduced whose solution defines the local randomized policy. The balancing authority broadcasts a common real-time control signal to the loads; at each time, each load changes state based on its own current state and the value of the common control signal.

2. The mean field limit defines an aggregate model for grid-level control. Special structure of the Markov model leads to a simple linear time-invariant (LTI) approximation. The LTI model is passive when the nominal Markov model is reversible.

3. Additional local control is used to put strict bounds on individual quality of service of each load, without impacting the quality of grid-level ancillary service.

Examples of application include chillers, flexible manufacturing, and even residential pool pumps. Sean demonstrated, through simulation, how pool pumps in Florida can supply a substantial amount of the ancillary service needs of the Eastern U.S.

Overview talk – VENKAT ANANTHARAM: *The objective method as a tool in large systems.* The objective method, also called the method of local weak convergence, is a technique to describe the local experience of a typical agent in a large interacting system through the solution of a recursive equation. Unlike the popular mean field limit method, it allows one to retain the essence of the locality of the interaction instead of modeling it away. This tool has so far been largely developed by probabilists and mathematicians in their quest to understand long range interactions in models motivated by physics and computer science. This tool appears to be potentially also of great value for understanding the role of interactions in control,communications, and other system-theoretic applications. Venkat provided an introduction to this technique, with illustrative examples revolving around distributed resource allocation.

ROLAND MALHAME: *A class of interference induced games: Asymptotic Nash equilibria and parameterized cooperative solutions.* Roland's talk focused on a multi-agent systems with linear stochastic individual dynamics and individual linear quadratic ergodic cost functions. The individuals partially observe their own states. Their cost functions and initial statistics are a priori independent, but they are coupled through an interference term (the mean of all agents' states), entering each of their individual measurement equations. While in general for a finite number of agents, the resulting optimal control law may be a non-linear function of the available observations, Roland showed that, for certain classes of costs and dynamic parameters, optimal separated control laws obtained by ignoring the interference coupling are asymptotically optimal when the number of agents goes to infinity, More generally though, optimal separated control laws may not be asymptotically optimal, and can in fact result in unstable overall behavior. Roland described a way of mitigating this via a class of parameterized decentralized control laws whereby the separated Kalman gain is treated as the arbitrary gain of a Luenberger-like observer. He characterized system stability regions and the nature of optimal cooperative control policies within the considered class, and presented some results of numerical simulations.

BILL SANDHOLM: *Large deviations and stochastic stability in the small noise double limit.* Bill Sandholm's talk was also on the subject of multi-agent games, but from the viewpoint of evolutionary dynamics. He described a model of stochastic evolution under general noisy best-response protocols, allowing the probabilities of suboptimal choices to depend on their payoff consequences. The focus of his talk was on behavior in the small-noise double limit: we first take the noise level in agents' decisions to zero, and then take the population size to infinity. He showed that, in this double limit, escape from (and transitions between) equilibria can be described in terms of solutions to certain continuous optimal control problems. These were used in turn to characterize the asymptotics of the stationary distribution, and thus to determine the stochastically stable states. He then presented a class of examples, where the control problems can be solved explicitly – three-strategy coordination games that satisfy the marginal bandwagon property and that have an interior equilibrium, and choices governed by the logit rule. He gave evidence for the conjecture that this should remain true for other classes of games and other choice rules.

PRAKASH NARAYAN: *Interactive multiterminal communication.* Shifting gears a bit, Prakash presented an information-theoretic perspective on communication in multi-agent systems. Information-theoretic models for multiuser source and channel coding usually take the communication between multiple terminals to be “simple,” i.e., noninteractive. On the other hand, studies of multiparty function computation, especially in computer science, emphasize the useful role of interactive communication. Prakash described basic structural properties of interactive communication and presented “Single-shot” bounds for the amount of common randomness, i.e., shared information, that can be generated among the terminals using such communication.

2.2 Day 2: Decentralized systems with partial observations, sequential decompositions, and sufficient statistics.

The first three talks, starting with an overview talk by Shlomo Zilberstein and followed by Matthijs Spaan and Frans Oliehoek, focused on decentralized partially observable Markov decision processes (DEC-POMDPs) from a computer science and artificial intelligence (AI) perspective.

Overview talk – SHLOMO ZILBERSTEIN: *Cooperative multi-agent planning in partially observable uncertain worlds.* Coordinating the operation of a group of decision makers or agents in stochastic environments is a long-standing challenge in AI. Decision theory offers a normative framework for optimizing decisions under uncertainty, but due to computational complexity barriers, developing decision-theoretic planning algorithms for multiagent systems has been a serious challenge. Shlomo described a range of recently developed formal models and algorithms to tackle this problem, largely based on extensions of MDPs and POMDPs. Exact algorithms shed light on the structure and complexity of the problem and offer valuable building blocks for approximation methods. He described a number of effective approximation techniques that use bounded memory, sampling, and randomization. These methods can produce high-quality results in a variety of application domains such as mobile robot coordination and sensor network management. He also examined the performance of these algorithms and describe current research efforts to further improve their applicability and scalability.

MATTHIJS SPAAN: *Multi-agent planning under uncertainty with communication.* Next, Matthijs presented decision-theoretic approaches to multiagent planning under uncertainty, formalized as extensions of the POMDP model. In cooperative settings, communication between agents has the potential to significantly improve team performance. For instance, a higher degree of coordination can often be obtained by sharing local information. A common objective, however, is to minimize the level of communication required for satisfying performance. Matthijs discussed different models of communication that have been explored and paired them with appropriate solution techniques. A particular focus was on multi-robot planning problems, in which quality restrictions on the communication network need to be taken into account.

FRANS OLIEHOEK: *Exploiting structured representation in multi-agent sequential decision making.* While the framework of DEC-POMDPs gives a principled formalization of multi-agent decision-making problems, computing optimal plans for them is computationally intractable (NEXP-complete). In order to try and scale-decision making to larger problems involving more agents, researchers have tried to exploit different forms of structure which may be present across many mult-iagent applications. Frans surveyed two types of approaches to exploiting structure in mult-iagent decision problems. The first type was based on methods from (distributed) constraint optimization. Essentially, these methods try to exploit conditional independence between agents. The second type was based on the insight that, even in the case of two agents that directly influence each other, not all details of the policies may be relevant for the coordination process. That is, rather than coordinating at the level of (joint) policies, it may be possible to coordinate at the more abstract level of (joint) influences of such policies.

Overview talk – MIHAILO JOVANOVIC: *Sparsity-promoting optimal control of distributed systems.* Mihailo’s talk, delivered from the perspective of a control theorist, was about the design of feedback gains that achieve a desirable tradeoff between quadratic performance of distributed systems and controller sparsity. (Similar ideas were explored by Ali Belabbas during Day 1.) He first identified sparsity patterns of the feedback gains by incorporating sparsity-promoting penalty functions into the optimal control problem, where the added terms penalize the number of communication links in the distributed controller. He then showed how to optimize feedback gains subject to structural constraints determined by the identified sparsity patterns. In the first step, the sparsity structure of feedback gains is identified using the alternating direction method of multipliers, an algorithm well-suited to large optimization problems. This method alternates between promoting the sparsity of the controller and optimizing the closed-loop performance, which allows us to exploit the structure of the corresponding objective functions. Even though the standard quadratic performance index is in general a nonconvex function of the feedback gain, we identify several classes of convex problems. In this case, the corresponding synthesis problem can be formulated as a semidefinite program for which we develop efficient customized solvers. Several examples were provided to demonstrate the effectiveness of the developed approach.

WING S. WONG: *Systems for open interaction.* For the majority of distributed multi-agent systems reported in the literature the system objective is uniquely defined. However, there are many real-life examples in which the system objective could depend on the choices selected by the agents that are unknown to each other. Wing's talk designated these as systems for open interaction. Control design of such systems does not follow the classic paradigm, and is intimately tied with communication issues. Wing reported results on optimal control design for some concrete systems and used them to illustrate a number of fundamental issues of interest for cooperative decentralized systems.

RAHUL JAIN: *The art of sequential optimization via simulations.* Sequential optimization is a common thread that emerges in the context of decentralized decision-making. Rahul presented a natural framework for simulation-based optimization and control of MDPs. The idea is very simple: Replace the Bellman operator that appears in the dynamic programming recursion by its 'empirical' variant wherein the expectation is replaced by a sample average approximation. This leads to a *random* Bellman operator. Rahul introduced several notions of probabilistic fixed points of such random operators, and showed their asymptotic equivalence. He then established convergence of empirical Value and Policy Iteration algorithms by a stochastic dominance argument. The mathematical technique introduced by Rahul is useful for analyzing other iterated random operators (and not just the empirical Bellman operator), and may also be useful in random matrix theory. The idea can be generalized to asynchronous dynamic programming, and is also useful for computing equilibria of zero-sum stochastic games. Preliminary numerical results showed better convergence rate and runtime performance than stochastic approximation/reinforcement learning, or any other commonly used schemes.

ADITYA MAHAJAN: *The common information approach to decentralized stochastic control.* Many modern technological systems, such as cyber-physical systems, communication, transportation and social networks, smart grids, sensing and surveillance systems are informationally decentralized. A key feature of informationally decentralized systems is that decisions are made by multiple decision makers that have access to different information. This feature violates the fundamental assumption upon which centralized stochastic control theory is based, namely, that all decisions are made by a centralized decision maker who has access to all the information and perfectly recalls all past observations and decisions/actions. Consequently, techniques from centralized stochastic control cannot be directly applied to decentralized stochastic control problem primarily for the following reason. In centralized stochastic control, the controllers belief on the current state of the system is a *sufficient statistic* for decision making. A similar sufficient statistic does not work for decentralized stochastic control because controllers have different information and hence their beliefs on the state of the system are not consistent. Nevertheless, two general approaches that use ideas from centralized stochastic control theory have been used for the solution of decentralized control problems: (i) the person-by-person approach; and (ii) the designers approach. Aditya gave a detailed discussion of the features and merits of these approaches, as well as some illustrative examples, such as control problems with delay-sharing, control-sharing, and mean-field information sharing.

SANJAY LALL: *Sufficient statistics for multi-agent decisions.* Continuing in the same vein, Sanjay described a notion of sufficient statistics for decision, estimation or control problems involving multiple players. As in the classical single-player setting, sufficient statistics contain all of the information necessary for the players to make optimal decisions. In the multi-agent setting, he showed how to construct such sufficient statistics via a convex relaxation of the feasible set of the corresponding decision problem. He proved that that these statistics may be updated recursively, and may be constructed by appropriately composing the corresponding single-player statistics. He also presented algorithms for this construction when the information pattern is defined by an appropriate graph.

NUNO MARTINS: *Distributed estimation over shared networks: optimal event-based policies.* Nuno focused on the design of distributed estimation systems that are composed of multiple non-collocated components. A shared network is used to disseminate information among the components. He discussed recent results that characterize the structure of certain optimal policies for the case in which the number of components exceeds the maximal number of simultaneous transmissions that the network can accept. In order to obtain a tractable framework for which design principles can be characterized analytically, he consider at most two estimators that rely on information sent to them by two sensors that access dissimilar measurements. He showed the optimality of certain threshold-based policies, established a connection with a problem of optimal quantization for which the distortion is non-uniform across representation symbols, presented numerical approaches, discussed interpretations of the results, and listed related open issues.

BAHMAN GHARESIFARD: *On the convergence of strategic interaction dynamics on networks.* Bahman presented a proof of the following result: piecewise-linear best-response dynamical systems with strategic interactions are asymptotically convergent to their set of equilibria on any weighted undirected graph. He discussed various features of these dynamical systems, including the uniqueness and abundance properties of the set of equilibria and the emergence of unstable equilibria. He also introduced the novel notions of *social equivalence* and *social dominance* for such dynamics on directed graphs, and demonstrated some of their interesting implications, including their correspondence to consensus and chromatic number of partite graphs.

2.3 Day 3: Information and learning in multi-agent cooperation and competition.

Overview talk – JASON MARDEN: *The role of information in multi-agent coordination.* The goal in networked control of multi-agent systems is to induce desirable collective behavior through the design of local control algorithms. The information available to the individual agents, either attained through communication or sensing, invariably defines the space of admissible control laws. Hence, informational restrictions impose constraints on achievable performance guarantees. Jason talked about one such constraint with regards to the efficiency of the resulting stable solutions for a class of networked resource allocation problems with submodular objective functions. When the agents have full information regarding the mission space, the efficiency of the resulting stable solutions is guaranteed to be within 50% of optimal. However, when the agents have only localized information about the mission space, which is a common feature of many well-studied control designs, the efficiency of the resulting stable solutions can be $1/n$ of optimal, where n is the number of agents. Consequently, in general such schemes cannot guarantee that systems consisting of n agents can perform any better than a system consisting of a single agent for identical environmental conditions. Jason presented an algorithm that overcomes this limitation by allowing the agents to communicate minimally with neighboring agents.

GÜRDAL ARSLAN: *Learning algorithms for stochastic dynamic games.* Gürdal first provided an overview of the literature on learning in static and dynamic games and made the point that there do not exist decentralized learning algorithms for dynamic games which are uncoupled (that is, which do not require the decision maker's to share any information) which converge to equilibria. Generalizing the notion of *weakly acyclic games* for static games to dynamic games, he introduced a decentralized Q-learning algorithm which is guaranteed, first probabilistically and then (through annealing) almost surely, to equilibria. Such games include dynamic team problems and potential games.

DAVID LESLIE: *Two-timescales game-theoretical learning with continuous action spaces.* Much work on learning in games has focused on the discrete-action case. However, in the contexts of both economics and control applications, the action set is frequently continuous (such as the price of a good, the position, or power level of a component of the system). David talked about a framework for learning in games with continuous actions sets, and introduced the necessary stochastic approximation theory with which such processes can be analyzed. He then demonstrated how this theory can be extended to a two-timescales system, in which the values of actions can be estimated, and strategies adapted, without any player actually observing the actions or rewards of any other. This results in a system where 'individual action learners' can converge successfully to approximate Nash equilibria, in zero-sum games and potential games.

JOHANNES HÖRNER: *Dynamic Bayesian games.* Johannes focused on characterizing an equilibrium payoff subset for Markovian games with private information as in the limit of vanishing discount factor. Monitoring is imperfect, transitions may depend on actions, types may be correlated and values may be interdependent. The focus was on equilibria in which players report truthfully. This characterization generalizes that for repeated games, reducing the analysis to static Bayesian games with transfers. With correlated types, results from mechanism design apply, yielding a folk theorem (or general feasibility theorem). With independent private values, the restriction to truthful equilibria is without loss, except for the punishment level; if players withhold their information during punishment-like phases, a "folk theorem" obtains also.)

BEHROUZ TOURI: *TS graph approach to evolutionary stability.* Behrouz revisited the theory of evolutionary stability for population games in finite populations. He provided two illustrative examples showing that (i) games with no evolutionarily stable strategy can admit a cloud of dominating set of strategies emerging from the selection-mutation process, and (ii) there are games with a unique evolutionarily stable strategy which, interestingly, becomes extinct in the selection-mutation dynamics. He generalized the conventional theory of evolutionary stability and showed how a cloud of strategies can emerge in the selection-mutation dynamics. The key object in our development is the Transitive Stability (TS) graph of the population dynamics. Behrouz demonstrated that all the strategies, except those of the extreme vertices of the associated TS-graph, become extinct in the mutation-selection process.

ALEX OLSHEVSKY: *Linear-time consensus.* Alex described a protocol for the average consensus problem on any fixed undirected graph whose convergence time scales linearly in the total number nodes n. The protocol is completely distributed, with the exception of requiring each node to know an upper bound on the total number of nodes which is correct within a constant multiplicative factor. He discussed applications of this protocol to questions in decentralized optimization and multi-agent control connected to the consensus problem. In particular, we develop a distributed protocol for minimizing an average of (possibly nondifferentiable) convex functions. Under the same assumption as above, and additionally assuming that the subgradients of each term in the average have norms upper-bounded by some constant L known to the nodes, after T iterations our protocol has error which is $O(L\sqrt{n/T})$.

SHREYAS SUNDARAM: *Spectral and structural properties of random graphs with applications to consensus dynamics on networks.* Shreyas' talk focused on some recent results on spectral and structural properties of large-scale random graphs that play a role in certain classes of coordination and consensus dynamics on networks. He started with a commonly studied class of Laplacian dynamics where the asymptotic rate of convergence is given by the second smallest eigenvalue (also known as the algebraic connectivity) of the graph Laplacian matrix. We provide a characterization of the algebraic connectivity of certain random graphs,including some results pertaining to interdependent networks that highlight the relative importance of edges between and within communities in the network. We then consider a variant of Laplacian dynamics where one or more nodes in the network keep their state fixed at certain values – such scenarios occur in the case of diffusion dynamics with stubborn agents.The rate of convergence in such settings is given by the smallest eigenvalue of the grounded Laplacian matrix of the network; we give explicit characterizations of this eigenvalue for various random graphs. Finally, we study a class of dynamics where each node in the network ignores the most extreme values in its neighborhood before averaging the rest, and propose a topological property known as "robustness" to characterize networks where such dynamics will lead to consensus. We show that this notion of robustness is much stronger than classical metrics such as connectivity, but that these properties coincide in certain random graph models (Erdős-Renyi, 1-D geometric, and preferential attachment graphs).

MAXIM RAGINSKY: *Online discrete optimization in social networks with inertia.* Recently, there has been a great deal of interest in modeling and analysis of collective information processing by networks of locally interacting entities with limited information. On the theoretical front, this circle of problems has deep connections to ideas from game theory, statistical physics, and discrete probability. Maxim described a model of online (i.e., real-time) discrete optimization by a social network consisting of agents that must choose actions to balance immediate time-varying costs against a desire to act according to some default myopic strategy. The costs are generated by a dynamic environment, and the agents lack ability or incentive to construct an a priori model of the environment's evolution. The global cost of the network decomposes into a sum of individual and pairwise interaction terms, and at each time step each agent is informed only about its own cost and the pairwise costs in its immediate neighborhood. The overall objective is to minimize the worst-case regret, i.e., the difference between the cumulative real-time performance of the network and the best performance that could have been achieved in hindsight with full centralized knowledge. Maxim presented an explicit strategy for the network based on the Glauber dynamics and showed that it achieves favorable scaling of the regret in terms of problem parameters under a Dobrushin-type mixing condition. The proof relied on ideas from statistical physics, as well as on recent developments in the theory of Markov chains in metric spaces, specifically Yann Ollivier's notion of positive Ricci curvature of a Markov operator.

ANDRAS GYÖRGY: *Online learning in Markov Decision Processes with changing reward sequences.* Andras talked about consider online learning in finite stochastic Markovian environments, where at each time step a new reward function is chosen by an oblivious adversary. The goal of the learning agent is to compete with the best stationary policy in hindsight in terms of the total reward received. Two variants of the problem were considered: the online stochastic shortest path problem and learning in unichain Markov decision processes. Several low-complexity algorithms were proposed, for both the full-information and the bandit-feedback settings, which achieve almost optimal performance under different assumptions on the Markov transition kernel. In the case of full-information feedback, the results presented by Andras complement existing results, while in the bandit feedback he gave the first low-complexity algorithms achieving optimal performance. Some of the algorithms were based on distributed learning, where separate learning agents were used for each state of the MDP, while others were completely centralized. Andras compared these approaches and discussed how different levels of synchronization and information sharing affects performance.

CSABA SZEPESVARI: *Online learning under delayed feedback.* In distributed, networked systems, feedback arrives in a delayed fashion. The question then arises how this will affect the learning speed of various algorithms. Csaba talked about some recent results concerning this effect in the framework of online learning with partial monitoring, which includes both full-information feedback, as well as bandit information feedback, as special cases. He then described novel results concerning the impact of delay on the regret, both for adversarial and stochastic settings. These results showed that, while for the adversarial setting the cost of the delay increases the regret in a multiplicative way, in the stochastic setting the increase is only additive. He described meta-algorithms that transform, in a black-box fashion, algorithms developed for the non-delayed case into ones that can handle the presence of delays in the feedback loop. Modifications of the well-known UCB algorithm were also developed for the stochastic multi-armed bandit problem with delayed feedback, with the advantage over the meta-algorithms that they can be implemented with lower complexity. Remaining open problems are discussed at the end.

VIANNEY PERCHET: *Blackwell approachability in (stochastic) games.* Vianney gave a unified presentation of standard frameworks for Blackwell approachability, both in full and partial monitoring, by defining a new abstract game, called the purely informative game, where the outcome at each stage is the maximal information players can obtain, represented as some probability measure. Objectives of players can be rewritten as the convergence (to some given set) of sequences of averages of these probability measures. He presented new results extending the approachability theory developed by Blackwell. This new abstract framework enables us to characterize approachable sets with a remarkably simple and clear reformulation for convex sets. Translated into the original games, those results become the first necessary and sufficient condition under which an arbitrary set is approachable and they cover and extend previous known results for convex sets. He also described a specific class of games where, thanks to some unusual definition of averages and convexity, one can obtain a complete characterization of approachable sets along with rates of convergence.

2.4 Day 4. Information and communication constraints in decentralized systems.

Overview talk – SERDAR YÜKSEL: *Quantization for stabilization, optimization and approximation in networked control.* In this three-part talk, Serdar discussed three ways in which quantization appears in stochastic control and networked control. The first area is on stabilization under information constraints, for which quantization is essential. For non-linear stochastic systems, he presented lower and upper bounds on information transmission rates required for stabilization in the sense of asymptotic stationarity, which in the case of linear systems turn out to be equal, and established the connection with a well-known limitation theorem due to Bode interpreted for non-linear systems. The second area is on optimization under quantization constraints, where he presented structural and existence results for optimal quantization policies. Explicit solutions were obtained for the application of these results to Linear Quadratic Gaussian systems. Finally, in the third part, he presented the problem of approximation in stochastic and networked control, where the action sets available to a controller are restricted to be finite. Asymptotic optimality of controllers which only can select from finitely many actions, as the size of the sets increase to infinity, was established, and under further conditions on the controlled Markov chains, explicit rates of convergence bounds were established. These bounds are tight in the sense of *order-optimality* for a class of systems and cost functions.

ANANT SAHAI: *Information theory and decentralized control.* Networked control systems require an interdisciplinary study of information, control and probability. In his talk, Anant discussed the interaction of information and control theories for networked control systems. Viewing the control plant as the source to a noisy channel, he presented a novel look at arriving at the requirements on channel reliabilities for stochastic stabilization. He then studied erasure channels extensively and discussed the limitations in imperfections in the sensor-controller channel and the controller-plant channel, and showed that the optimal networked control problems are not robust to uncertainties in the information channel.

CÉDRIC LANGBORT: *On myopic strategies in dynamic adversarial team decision problems.* Cedric discussed game theoretic aspects in networked control and dynamic team problems, where there are decision makers with mis-aligned objective functions. In the context of a large class of cost functions, including the celebrated Witsenhausen's counterexample problem, the quadratic source-channel coding problem as well as variations of such quadratic cost functions, he discussed the optimality and sub-optimality of linear policies, and quantization based strategies for such problems. When uncertainty in the system model is included, he made the point that for a class of team/game problems, when the uncertainty as seen by a decision maker is quantitatively large, myopic strategies (non-strategic policies which are oblivious to other decision makers) may perform very well and be indeed optimal for a class of quadratic problems.

OHAD SHAMIR: *Information constraints in distributed and online learning.* Many machine learning approaches are characterized by information constraints on how they interact with training data. One prominent example is communication constraints in distributed learning, while other examples include memory and sequential access constraints in stochastic optimization and online learning, and partial access to the underlying data (such as in multi-armed bandits). However, in a statistical learning setting, we have little understanding how these information constraints fundamentally affect the learning performance, without making assumptions on their semantics. Ohad discussed these issues, and described how a single set of results sheds some light on them, simultaneously for different types of information constraints and for different learning settings.

VIVEK BORKAR: *Gossip and related algorithms.* Vivek concluded the workshop by describing some of his recent work on gossip and related distributed algorithms, highlighting some possible future directions.

3 Scientific Progress Made and Outcome of the Meeting

The main outcome of the workshop workshop was to provide an opportunity for researchers from different communities (stochastic control, economics, information theory, and machine learning) working on optimal decentralized decision-making to exchange ideas and learn new mathematical tools and techniques (used by other communities). This was an intense week, and each day was packed with many talks looking at a wide variety of decentralized decision-making problems using a broad array of tools from probability theory, information theory, systems and control, game theory, optimization, and artificial intelligence. Nevertheless, there were plenty of discussions among the participants during the coffee breaks and after dinner, and through many email communications after the workshop, and we believe that ultimately the workshop succeeded in bringing to the fore various connections between the solution approaches of the different communities, fostering future collaborations, and providing an improved understanding of optimal decentralized decision-making. Since this was the first workshop of this kind on the topic of decentralized decision-making, we expect to see the fruits of collaboration between the participants in the near future; in the meantime, here are some impressions of the workshop from several participants:

Excellent workshop, idyllic surroundings. Feel recharged.

Vivek Borkar
Electrical Engineering, Indian Institute of Technology Bombay

Thanks a lot for the invitation to Banff. I learnt a great deal, thanks to you. Hopefully we will get more chances to interact, stay in touch.

Johannes Hörner
Economics, Yale University

The workshop was well organized, gathering scholars from different disciplines, which opened new angles for me.

Michel De Lara
CERMICS, Ecole des Ponts ParisTech

The BIRS workshop was an outstanding experience. The opportunity to interact with and learn from the top researchers in the field, against the idyllic backdrop of the Canadian Rockies, was very memorable. I am very glad that places such as BIRS and the Banff Centre exist to facilitate the exchange and pursuit of knowledge.

Shreyas Sundaram
Electrical and Computer Engineering, University of Waterloo