



Operations Research

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

In This Issue

To cite this article:

(2022) In This Issue. Operations Research 70(6):iii-vii. <https://doi.org/10.1287/opre.2022.2411>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2022, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

In This Issue

<https://doi.org/10.1287/opre.2022.2411>

Copyright: © 2022 INFORMS

Endogenization of the Reference Point Reduces Loss Aversion

A central idea in behavioral economics is that agents derive utility from gain and losses relative to a certain reference point and that losses loom larger than gains. In “How Endogenization of the Reference Point Affects Loss Aversion: A Study of Portfolio Selection,” He and Strub study the implications of various models of partially endogenous reference points on portfolio selection. In these models, an agent faces a salient exogenous reference point that influences the formation of endogenously determined expectations about the future, through rational expectations, optimal expectations, or mental updating of the reference point. A key finding is that the predictions of these models are identical to a model with an exogenous reference point but a lower degree of loss aversion. This suggests that it is difficult to separately identify an agent’s degree of loss aversion and his or her reference point and may help to explain why experienced and sophisticated agents appear to be less loss averse than expected in some field settings.

Integrating Cost Reduction and Supplier Selection Activities

Manufacturers who outsource production may find ways to reduce suppliers’ cost to lower the contract price. Conventionally, this is done after supplier selection. However, in practice, a sizable portion of the degree to which the manufacturer can help a supplier reduce cost could be learned before the supplier starts production. This opens up a possibility to better integrate supplier selection and cost-reduction activities. In “Procurement Mechanisms with Postauction Preaward Cost-Reduction Investigations,” Chen, Beil, and Duenyas explore an alternative procurement process wherein, after suppliers’ initial pricing bids are collected, the manufacturer investigates suppliers’ cost-saving potential and then leverages this information to select the contract winner. They provide analytical and numerical evidence that this integrated procurement process not only helps the manufacturers reduce cost significantly, but also benefits the suppliers in many plausible scenarios.

Continuous Patrolling

Sometimes it is necessary to have a Patroller (on foot, in a car, or maybe a drone) move around a network so as to prevent an intruder (the Attacker) from infiltrating or otherwise ruining the network operation. The “attack” could be, for example, removing a painting from the Louvre, crossing a border, or planting a bomb. The first possibility could take place at only a discrete set of points on the network, say, nodes. However, the last two types of attack could take place anywhere. The latter continuous problem has been modelled as a game by Alpern, Bui, Lidbetter, and Papadaki in the article “Continuous Patrolling Games.” The Attacker decides when and where to attack (the duration of the attack is specified by the problem), whereas the Patroller chooses a unit speed path, possibly periodic. If the Patroller passes the attacked point while the attack is going on, he wins the game, as the attack

is thwarted. Otherwise, the attack is successful, and the Attacker wins the game. The authors determine optimal strategies for both players for many classes of networks and find good strategies that work on any network. These ideas could be adapted to real-life patrolling problems on networks.

Finding Social Media Users in a Location

In events such as natural disasters, terrorist attacks, or war zones, one can gain critical situational awareness by monitoring what people on the ground are saying in social media. But how does one build a set of users in a specific location from scratch? In “Building a Location-Based Set of Social Media Users,” Marks and Zaman present an algorithm to do just this. The algorithm starts with a small set of seed users in the location and then grows this set using an “expand-classify” approach. They apply the algorithm to diverse regions ranging from South America to the Philippines and in a few hours can collect tens of thousands of Twitter users in the target locations. The algorithm is language agnostic, making it especially useful for anyone trying to gain situational awareness in foreign countries.

Dark Matter in (Volatility and) Equity Option Risk Premiums

In “Dark Matter in (Volatility and) Equity Option Risk Premiums,” Bakshi, Crosby, and Gao ask a provocative question: Is there dark matter embedded in volatility and equity options? They consider a theoretical approach that allows them to introduce the constructs of risk premiums on jumps crossing the strike and on local time. The treatment of jumps crossing the strike and local time is integral to their theory because their absence would be counterfactual from an empirical standpoint. They label such abstract uncertainties—driven by unspanned risk components—“dark matter” as these uncertainties can be hard to identify, but their presence is implied in options data, and the workings of dark matter can be economically influential. Their empirical exercises are based on weekly equity index options (the “weeklys”) in addition to the farther dated (index and futures) options up to 88 days maturity.

How Should Fashion Brands Use Advertising to Increase Sales While Remaining Exclusive?

Fashion consumption signals a consumer’s status to the broader population, so fashion brands and their consumers value exclusivity. Consequently, fashion advertising must balance sales generation with exclusivity loss. In “Advertising Cycling to Manage Exclusivity Loss in Fashion Styles,” Bruce, Krishnamoorthy, and Prasad develop a model with these features of fashion and estimate it using advertising, price, and sales data for two styles of handbags and sunglasses. Their analysis provides insights for advertising budgeting and scheduling and finds that advertising optimally should decrease as the product increases in popularity and vice versa. This exerts a braking force on sales oscillations so that the fashion cycle decays as does the optimal advertising

path. In addition to demonstrating how advertising cycling can impact a fashion firm's profitability, they show how different styles of a fashion brand can cycle at different rates. By connecting advertising cycles to fashion cycles, they provide prescriptions for how fashion firms should manage different styles of the same brand.

Entrywise Bounds for Sparse Matrix Estimation via Collaborative Filtering

Matrix estimation or completion has served as a canonical mathematical model for recommendation systems. More recently, it has emerged as a fundamental building block for data analysis as a first step to denoise the observations and predict missing values. Since the dawn of e-commerce, similarity-based collaborative filtering has been used as a heuristic for matrix estimation. At its core, it encodes typical human behavior: you ask your *friends* to recommend what you may *like* or *dislike*. Algorithmically, *friends* are *similar* “rows” or “columns” of the underlying matrix. The traditional heuristic for computing similarities between rows has costly requirements on the density of observed entries. In “Iterative Collaborative Filtering for Sparse Matrix Estimation” Borgs, Chayes, Shah, and Yu introduce an algorithm that computes similarities in *sparse* datasets by comparing expanded local neighborhoods in the associated data graph: in effect, you ask friends of your friends to recommend what you may like or dislike. This work provides bounds on the max entry-wise error of their estimate for low rank and approximately low rank matrices, which is stronger than the aggregate mean squared error bounds found in classical works. The algorithm is also interpretable, scalable, and amenable to distributed implementation.

Making Data-Driven Estimation Generalizable When Data Are Scarce

In “High-Dimensional Learning Under Approximate Sparsity with Applications to Nonsmooth Estimation and Regularized Neural Networks,” Liu, Ye, and Lee study a model fitting problem where there are much fewer data than problem dimensions. Of their particular focus are the scenarios where the commonly imposed sparsity assumption is relaxed, and the usual condition of the restricted strong convexity is absent. The results show that generalization performance can still be ensured in such settings, even if the problem dimensions grow exponentially. The authors further study the sample complexities of high-dimensional nonsmooth estimation and neural networks. Particularly for the latter, it is shown that, with explicit regularization, a neural network is provably generalizable, even if the sample size is only polylogarithmic in the number of fitting parameters.

Adapting to Unknown Payoff Smoothness in Non-Parametric Contextual Bandits

In nonparametric contextual bandit formulations, a key complexity driver is the smoothness of payoff functions with respect to covariates. In many practical settings, the smoothness of payoffs is unknown, and misspecification of smoothness may severely deteriorate the performance of existing methods. In “Smoothness-Adaptive Contextual Bandits,” Gur, Momeni, and Wager consider a framework where the smoothness of payoff functions is unknown and study when and how algorithms may adapt to unknown smoothness. First, they establish that designing algorithms that adapt to unknown smoothness is, in general, impossible. However, under a natural self-similarity condition, they establish that adapting to unknown smoothness is possible and devise a general policy for achieving smoothness-adaptive performance. The policy infers the smoothness of payoffs throughout the decision-making process while leveraging the structure of off-

the-shelf nonadaptive policies. It matches (up to a logarithmic scale) the performance that is achievable when the smoothness of payoffs is known in advance.

Data Science for Motion and Time Analysis

Motion and time analysis has been a popular tool in operations research for analyzing work performance in manufacturing and service operations. The current practice in motion and time analysis involves many labor-intensive steps such as stop-watching, videotaping, and manual data analysis. It is too inefficient to be practiced regularly for continuous improvement. Whereas modern sensing devices have automated and eased motion measurements, the motion analytics transforming the new data into knowledge are largely underdeveloped. Unsolved technical questions include: How can the motion and time information be extracted from the motion sensor data? How are work motions and work rates statistically modeled and compared? How are the motions correlated to the rates? In “Data Science for Motion and Time Analysis with Modern Motion Sensor Data,” Park, Noh, and Srivastava develop solutions to the technical questions into a novel data science framework for motion and time analysis. The new framework is demonstrated with industrial use cases for a smart factory.

A Theoretical Understanding of Monte Carlo Tree Search

In “Non-Asymptotic Analysis of Monte Carlo Tree Search,” Shah, Xie and Xu consider the popular tree-based search strategy, the Monte Carlo Tree Search (MCTS), in the context of infinite-horizon discounted Markov Decision Process. They show that MCTS with an appropriate polynomial rather than logarithmic bonus term indeed leads to the desired convergence property. The authors derive the results by establishing polynomial concentration property of regret for a class of non-stationary Multi-Arm Bandits. Further, using this as a building block, they demonstrate that MCTS, combined with nearest neighbor supervised learning, acts as a “policy improvement” operator that can iteratively improve value function approximation.

Dynamic Personalized Decision Making Beyond the Super-Extrapolatable and Super-Local Cases

Contextual bandit problems model the inherent trade-off between exploration and exploitation in personalized decision making in marketing, healthcare, revenue management, and more. Specifically, the trade-off is characterized by the optimal growth rate of the regret. Intuitively, the optimal rate should depend on how complex the underlying supervised learning problem is, namely, how much can observing reward in one context tell us about mean rewards in another. To formalize this intuitive relationship, Hu, Kallus, and Mao study in “Smooth Contextual Bandits: Bridging the Parametric and Nondifferentiable Regimes” a nonparametric contextual bandit problem in which mean reward functions are β -times differentiable (more generally, Hölder β -smooth). This interpolates between two extremes previously studied in isolation: nondifferentiable bandits ($\beta \leq 1$), with which running separated noncontextual bandits in different context regions achieves rate-optimal regret, and parametric-response bandits ($\beta = \infty$), with which rate-optimal regret can be achieved with minimal or no exploration because of infinite extrapolatability across contexts. The authors develop a rate-optimal algorithm that operates neither fully locally nor fully globally, revealing the optimal regret rate in this in-between smooth setting and shedding light on the crucial interplay of functional complexity and regret in dynamic personalized decision making.

Demystifying the Curse of Horizon in Offline Reinforcement Learning in Order to Break It

Offline reinforcement learning (RL), where we evaluate and learn new policies using existing off-policy data, is crucial in applications where experimentation is challenging and simulation unreliable, such as medicine. It is also notoriously difficult because the similarity (density ratio) between observed trajectories and those generated by any new policy diminishes exponentially as the horizon grows, known as the curse of horizon, which severely limits the application of offline RL whenever horizons are moderately long or even infinite. In “Efficiently Breaking the Curse of Horizon in Off-Policy Evaluation with Double Reinforcement Learning,” Kallus and Uehara set out to understand these limits and when they can be broken. They precisely characterize the curse by deriving the semiparametric efficiency lower bounds for the policy-value estimation problem in different models. On the one hand, this shows why the curse necessarily plagues standard estimators: they work even in non-Markov models and therefore must be limited by the corresponding bound. On the other hand, greater efficiency is possible in certain Markovian models, and they give the first estimator achieving these much lower efficiency bounds in infinite-horizon Markov decision processes.

Integrating Predictions in CPLEX’s Algorithmic Design: MIQP Linearization

Despite modern solvers being able to tackle mixed-integer quadratic programming problems (MIQPs) for several years, the theoretical and computational implications of the employed resolution techniques are not fully grasped yet. An interesting question concerns the choice of whether to linearize the quadratic part of a convex MIQP: although in theory no approach dominates the other, the decision is typically performed during the preprocessing phase and can thus substantially condition the downstream performance of the solver. In “A Classifier to Decide on the Linearization of Mixed-Integer Quadratic Problems in CPLEX,” Bonami, Lodi, and Zarpellon use machine learning (ML) to cast a prediction on this algorithmic choice. The whole experimental framework aims at integrating optimization knowledge in the learning pipeline and contributes a general methodology for using ML in MIP technology. The workflow is fine-tuned to enable online predictions in the IBM-CPLEX solver ecosystem, and, as a practical result, a classifier deciding on MIQP linearization is successfully deployed in CPLEX 12.10.0.

Mixed-Projection Conic Optimization: A New Paradigm for Modeling Rank Constraints

Many central problems throughout optimization, machine learning, and statistics are equivalent to optimizing a low-rank matrix over a convex set. However, although rank constraints offer unparalleled modeling flexibility, no generic code currently solves these problems to certifiable optimality at even moderate sizes. Instead, low-rank optimization problems are solved via convex relaxations or heuristics that do not enjoy optimality guarantees. In “Mixed-Projection Conic Optimization: A New Paradigm for Modeling Rank Constraints,” Bertsimas, Cory-Wright, and Pauphilet propose a new approach for modeling and optimizing over rank constraints. They generalize mixed-integer optimization by replacing binary variables z that satisfy $z^2 = z$ with orthogonal projection matrices Y that satisfy $Y^2 = Y$. This approach offers the following contributions: First, it supplies certificates of (near) optimality for low-rank problems. Second, it demonstrates that some of the best ideas in mixed-integer optimization, such as decomposition methods, cutting planes, relaxations, and random rounding schemes, admit straightforward extensions to mixed-projection optimization.

When Service Times Depend on Customers’ Delays: A Relationship Between Two Models of Dependence

Service times of customers often depend on the delay they experience in queue, as was recently demonstrated empirically in restaurants, call centers, and intensive care units. Two forms of dependence mechanisms in service systems with customer abandonment are studied in this paper: First, the service requirement of a customer may evolve while waiting in queue. Second, customers may arrive to the system with an exogenous service and patience time that are stochastically dependent. Because either dependence mechanism can have significant impacts on a system’s performance, it should be identified and taken into consideration for performance evaluation and decision-making purposes. However, identifying the source of dependence from observed data is hard because both the service times and patience times are censored due to customer abandonment. Further, even if the dependence is known to be the latter exogenous one, there remains the difficult task of fitting a joint service-patience times distribution to the censored data. In “When Service Times Depend on Customers’ Delays: A Relationship Between Two Models of Dependence,” Wu, Bassamboo, and Perry provide a solution to address these statistical challenges.

Asymptotically Optimal Control of a Centralized Dynamic Matching Market with General Utilities

The utility of a match in a two-sided matching market often depends on a variety of characteristics of the two agents (e.g., a buyer and a seller) to be matched. In contrast to the matching market literature, this utility may best be modeled by a general matching utility distribution. In “Asymptotically Optimal Control of a Centralized Dynamic Matching Market with General Utilities,” Blanchet, Reiman, Shah, Wein, and Wu consider general matching utilities in the context of a centralized dynamic matching market. To analyze this difficult problem, they combine two asymptotic techniques: extreme value theory (and regularly varying functions) and fluid asymptotics of queueing systems. A key trade-off in this problem is market thickness: Do we myopically make the best match that is currently available, or do we allow the market to thicken in the hope of making a better match in the future while avoiding agent abandonment? Their asymptotic analysis derives quite explicit results for this problem and reveals how the optimal amount of market thickness increases with the right tail of the matching utility distribution and the amount of market imbalance. Their use of regularly varying functions also allows them to consider correlated matching utilities (e.g., buyers have positively correlated utilities with a given seller), which is ubiquitous in matching markets.

New Economic Characterization for Proper Weighted Scoring Rules

In “Weighted Scoring Rules and Convex Risk Measures,” Smith and Bickel present a general connection between weighted proper scoring rules and investment decisions involving the minimization of a convex risk measure. Weighted scoring rules are quantitative tools for evaluating the accuracy of probabilistic forecasts relative to a baseline distribution. The authors demonstrate that the relationship between convex risk measures and weighted scoring rules relates closely with previous economic characterizations of weighted scores based on expected utility maximization. As illustrative examples, the authors study two families of weighted scoring rules based on phi-divergences (generalizations of the Weighted Power and Weighted Pseudospherical Scoring rules) along with their corresponding risk measures. The paper will be of particular interest to the decision analysis and mathematical finance communities as well as those interested in the elicitation and evaluation of subjective probabilistic forecasts.

Characterizing Product-Form Solutions via Probabilistic Representations and Large Deviations

Single-class closed queueing networks, consisting of infinite-server and single-server queues with exponential service times and probabilistic routing, admit product-form solutions. Such solutions, although seemingly tractable, are difficult to characterize explicitly for practically relevant problems due to the exponential combinatorial complexity of its normalization constant (partition function). In “A Probabilistic Approach to Growth Networks,” Jelenković, Kondev, Mohapatra, and Momčilović develop a novel methodology, based on a probabilistic representation of product-form solutions and large-deviations concentration inequalities, which identifies distinct operating regimes and yields explicit expressions for the marginal distributions of queue lengths. From a methodological perspective, a fundamental feature of the proposed approach is that it provides exact results for order-one probabilities, even though the analysis involves large-deviations rate functions, which characterize only vanishing probabilities on a logarithmic scale.

How To Effectively Schedule Multi-Task Jobs in Modern Parallel-Computing Frameworks?

Jobs in modern parallel-computing frameworks, such as Hadoop and Spark, are subject to several constraints. In these frameworks, the data are typically distributed across a cluster of machines and is processed in multiple stages. Therefore, tasks that belong to the same stage (job) have a collective completion time that is determined by the slowest task in the collection. Furthermore, a task's processing time is machine dependent, and each machine is capable of processing multiple tasks at a time subject to its capacity. In “Scheduling Parallel-Task Jobs Subject to Packing and Placement Constraints,” by Shafiee and Ghaderi, multiple approximation algorithms with theoretical guarantees are provided to solve the problem under preemptive and nonpreemptive scenarios. The numerical results, using a real traffic trace, demonstrate that the algorithms yield significant gains over the prior approaches.

Naïve Learning in a Binary Action, Social Network Environment

In “Naïve Learning Through Probability Overmatching,” Arieli, Babichenko, and Mueller-Frank consider an environment where privately informed agents select a binary action repeatedly observing the past actions of their neighbors in a social network. Rational inference has been shown to be exceedingly complex in this environment. Instead, this paper focuses on boundedly rational agents that form beliefs according to discretized DeGroot updating and apply a decision rule that assigns a (mixed) action to each belief. It is shown that naïve learning, where the long run actions of all agents are optimal given their pooled private information, can be achieved in any strongly connected network if beliefs satisfy a high level of inertia and the decision rule coincides with probability overmatching. The main difference to existing naïve learning results is that here it is shown to hold (1) for binary rather than uncountable action spaces and (2) even for network and information structures where Bayesian agents fail to learn.

Optimistic Gittins Indices

In “Optimistic Gittins Indices,” Farias and Gutin propose a tightening sequence of optimistic approximations to the Gittins index. They show that the use of these approximations in concert with the use of an increasing discount factor appears to offer a compelling alternative to state-of-the-art index schemes proposed for the Bayesian multiarmed bandit problem. The authors prove that the use of these optimistic indices constitutes a regret optimal

algorithm. Perhaps more interestingly, the use of even the loosest of these approximations appears to offer substantial performance improvements over state-of-the-art alternatives while incurring little to no additional computational overhead relative to the simplest of these alternatives.

More Efficient Bayesian Optimization Through the Use of Common Random Numbers

Bayesian optimization is a powerful tool for expensive stochastic black-box optimization problems, such as simulation-based optimization or hyperparameter tuning in machine learning systems. In “Bayesian Optimization Allowing for Common Random Numbers,” Pearce, Poloczek, and Branke show how explicitly modeling the random seed in the Gaussian process surrogate model allows Bayesian optimization to exploit the structure in the noise and benefit from variance reduction provided by common random numbers. The proposed knowledge gradient with common random numbers acquisition function iteratively determines a combination of input and random seed to evaluate the objective. It automatically trades off reusing old seeds to benefit from the variance reduction through common random numbers and querying new seeds to avoid bias because of a small number of seeds. The proposed algorithm is analyzed theoretically and empirically shows superior performance compared with previous approaches on various test problems.

Simulation Solution Screening Using Functional Properties

Simulation models today give rise to problems with large numbers of simulated scenarios or solutions—more than can be simulated exhaustively. Fortunately, users of these models may be able to verify or infer properties, such as convexity, of a performance measure of interest when viewed as a function over the space of solutions. In “Plausible Screening Using Functional Properties for Simulations with Large Solution Spaces,” Eckman, Plumlee and Nelson introduce a framework in which such properties are exploited to avoid simulating solutions with unacceptable performances. Their methods solve optimization problems that measure how well the result of a limited simulation experiment agrees with the claim that a solution is acceptable. These methods deliver desirable statistical guarantees of confidence and consistency. Numerical experiments illustrate how functional properties coupled with small simulation experiments can avoid many simulations for simulation-optimization problems.

Learning and Control for an Information Provider

Much of the past learning and control literature has focused on the design of information acquisition processes for the demand side of information and has assumed that information supply is always genuine. However, in many economic and management settings, the information provider has incentives to strategically disseminate his/her private information, even in a possibly biased way. For example, a company may advertise deceptively to sell its products. In “Learning Manipulation Through Information Dissemination,” Keppo, Kim, and Zhang take the perspective of an information provider and study the optimal manipulation of a learning process through the adaptive design of (mis)information. The authors explicitly characterize both the optimal manipulation policy and the learner's belief process under such manipulation. They also extend their analysis to social learners who rely on public reviews to resist manipulation and show that social learning indeed mitigates misinformation in the long run.

Shortfall Risk Models When Information on Loss Function Is Incomplete

Utility-based shortfall risk measures effectively capture a decision maker's risk attitude on tail losses. In "Shortfall Risk Models When Information on Loss Function Is Incomplete," Delage, Guo and Xu consider a situation where the decision maker's risk attitude toward tail losses is ambiguous and introduce a robust version of shortfall risk, which mitigates the risk arising from such ambiguity. Specifically, they use some available partial information or subjective judgement to construct a set of plausible utility-based shortfall risk measures and define a so-called preference robust shortfall risk as through the worst risk that can be measured in this (ambiguity) set. The authors then apply the robust shortfall risk paradigm to optimal decision-making problems and demonstrate how the latter can be reformulated as tractable convex programs when the underlying exogenous uncertainty is discretely distributed.

Simulation Optimization Using an Actor-Critic Structure

Many systems arising in applications from engineering design, manufacturing, and healthcare require the use of simulation optimization (SO) techniques to improve their performance. In "Actor-Critic-Like Stochastic Adaptive Search for Continuous Simulation Optimization," Zhang and Hu propose a randomized approach that integrates ideas from actor-critic reinforcement learning within a class of adaptive search algorithms for solving SO problems. The approach fully retains the previous simulation data and incorporates them into an approximation architecture to exploit knowledge of the objective function in searching for improved solutions. The authors provide a finite-time analysis for the method when only a single simulation observation is collected at each iteration. The method works well on a diverse set of benchmark problems and has the potential to yield good performance for complex problems using expensive simulation experiments for performance evaluation.

Constant Regret Resolving Heuristics for Price-Based Revenue Management

Network revenue management (NRM) and its corresponding pricing question is one of the most fundamental problems in operations management. To alleviate the curse of dimensionality and the prohibitive cost of computing an exact solution using dynamic programming, computationally efficient resolving algorithms are proposed. The state-of-the-art analysis of the resolving heuristic establishes a logarithmic additive regret for price-based NRM problems. In "Constant Regret Resolving Heuristics for Price-Based Revenue Management," Y. Wang and H. Wang from the University of Florida and Georgia Institute of Technology, respectively, significantly advance the state-of-the-art analysis by showing a constant regret for resolving heuristics. Their theoretical advance is made possible by a novel, direct analysis of the exact DP solution.

Forecast Regularization Based on the Limits of Rational Disagreement

How much can rational people really disagree? If we can understand the limits of such disagreement, can we remove noise by labeling excess disagreement as irrational and then construct a group belief based on everyone's rational beliefs? In "Regularized Aggregation of One-Off Probability Predictions," Satopää proposes a Bayesian aggregator that requires no user intervention and can be computed efficiently even for a large number of one-off probability predictions. To illustrate, the aggregator is evaluated on predictions collected during a four-year forecasting tournament sponsored by the U.S. intelligence community. The aggregator improves the squared error (a.k.a., the Brier score) of simple averaging by around 20% and other commonly used aggregators by 10%–25%. This advantage stems almost exclusively from improved calibration. An R package called braggR implements the method and is available on CRAN.

ALSO-X and ALSO-X+: Better Convex Approximations for Chance Constrained Programs

A chance constrained program (CCP) aims to seek the best decision whose probability of violating the uncertainty constraints is within the prespecified risk level. As a CCP is often nonconvex and difficult to solve to optimality, much effort has been devoted to developing convex inner approximations for a CCP. The conditional value-at-risk (CVaR) has been known to be the best for more than a decade. In "ALSO-X and ALSO-X+: Better Convex Approximations for Chance Constrained Programs," Jiang and Xie study and generalize the ALSO-X, originally proposed by Ahmed, Luedtke, Song, and Xie (2017), for solving a CCP, and proves that ALSO-X is better than the state-of-the-art CVaR.

Scalable Reinforcement Learning for Multiagent Networked Systems

Highlighted by success stories like AlphaGo, reinforcement learning (RL) has emerged as a powerful tool for decision making in complex environments. However, the success of RL has thus far been limited to small-scale or single-agent systems. To apply RL to large-scale networked systems such as energy, transportation, and communication networks, a critical hurdle is the curse of dimensionality, because for these systems, the state and action space can be exponentially large in the number of nodes in the network. In "Scalable Reinforcement Learning for Multiagent Networked Systems," Qu, Wierman, and Li attempt to break this curse of dimensionality and designs a scalable RL method, named scalable actor critic (SAC), for large networked systems. The key technical contribution is to exploit the network structure to derive an exponential decay property, which enables the design of the SAC approach.