Judge's Commentary: The Outstanding Coal-Tipple Operations Papers

Jonathan P. Caulkins
Heinz School of Public Policy and Management
Carnegie Mellon University
Pittsburgh, PA 15213–3890

This problem is deceivingly difficult because it has characteristics that are familiar (queueing, inventory, scheduling) but does not fit neatly into any standard class of problems. As a result, even though teams tried methods as varied as linear programming and simulated annealing, essentially every team also used Monte Carlo simulation.

Unfortunately, many teams plunged too quickly into simulation and neglected to perform supporting analyses. Indeed, several teams used Monte Carlo methods to estimate quantities, such as the expected value of the minimum of independent uniform random variables, that can easily be found exactly. These papers, even when the simulations were well constructed, typically generated more numbers than insight.

The better papers, including all three Outstanding papers, augmented simulations with other analysis. The U.S. Military Academy team produced upper and lower bounds on costs to verify that the results of its simulation were reasonable. When Thursday's operations spill over into Friday, demurrage costs on Friday may be affected. However, exact analysis of the costs of such disruptions is exceedingly difficult, and most teams simply ignored them. The University of Alaska Fairbanks team, in contrast, attempted to construct lower and upper bounds for those costs.

Better papers were distinguished also by a more mature treatment of the assumptions, including performing sensitivity analysis with respect to those assumptions. The Cornell University team stood out because they did not jump to the conclusion that two crews can fill the coal tipple twice as quickly as one. Using a simple graph, the team showed how demurrage costs vary with how much a second crew speeds up the loading process.

The three outstanding papers did not monopolize insightful analysis. For example, several teams gave careful discussions of whether or not to give priority to a high-capacity train. One even considered preemptive priorities and concluded that a standard train should preempt a high capacity train when the high-capacity train is less than about 20% full. There were also several excellent analyses of when to use one crew or two to refill the tipple,

as a function of the amount of coal in the tipple, the time of day, and the number of trains still to come that day.

The questions of priority and number of crews were approached at different levels of sophistication. Some teams recommended a reasonable course of action with little explanation. Such recommendations are of little value, because they must be taken on faith and do not bring out general properties. Better papers gave either insightful intuitive arguments that produced more understanding or else mathematical proofs that were more persuasive. The best papers gave both.

More generally, good papers had a sense of perspective. Weaker papers worried excessively about minutiae, such as computer roundoff error and the fact that random number generators do not produce truly random numbers. Better papers recognized the general structure of the problem and designed their approach around it (Cornell); neatly presented and evaluated alternative decision rules (University of Alaska Fairbanks); and distilled their results into a concise, well-written summary (U.S. Military Academy).

When addressing a real-world modeling problem, there is no such thing as having "finished" the problem: There are always more ways of interpreting, structuring, and approaching the problem. In fact, that is the value of publishing the Outstanding papers—reading them helps provoke contestants to think in new ways about the problem that they worked on. Several ideas that were not fully tried in any paper but occurred to the judges as potentially fruitful include explicitly defining a state-space description of the system, and using dynamic programming or Markov decision processes, either as a solution procedure or as a way of structuring the problem.

The judges rewarded teams who used multiple methods, conducted sensitivity analyses, derived and justified insightful properties of good solutions, and had the perspective to understand and acknowledge the limitations of their models. It is very difficult to do well in all of these dimensions; good modeling is an art that takes considerable skill and practice for proficiency. The judges are delighted that so many students have accepted the challenge to become good modelers.

About the Author

Jonathan P. Caulkins participated in the MCM at Washington University, where his teams' entries were judged Outstanding in the first two competitions. Jon earned a master's degree in electrical engineering and computer science and a doctorate in operations research from MIT. Now he is an assistant professor of operations research and public policy at Carnegie Mellon University's Heinz School of Public Policy and Management, where his research focuses on mathematical models of illicit drug markets. Jon was an associate judge for the Coal-Tipple Operations Problem.