Academy 32: Fireside Chat with Optimization Pioneer, Dr. Ed Rothberg

Question and Answer

We had a lot of great questions in follow-up to this Academy discussion. So many, in fact, that we didn't have time to answer them all live. Please see below for a list of all the questions and the corresponding answers from Mike and Ed. This includes both questions that were asked during the live session AND those we didn't have a chance to address at that time.

Question: It was mentioned in the chat that ML and optimization are somewhat mutually exclusive, i.e. one possibly feeds the other. However, is there a perception in the industry that ML/Data Science is more impactful than optimization? I never see as much excitement about optimization as there is for ML.

> **Answer (Ed):** Yes, it's definitely true. This is basically just the fact that optimization has been around for longer. At this point we know what it can do, there are always new things that are being discovered, but the edges are known. Whereas with Machine Learning it's a new technology. It's being applied in new ways so you don't necessarily know where the boundaries are. The claim I often make is that the impact of optimization (on business) absolutely dwarfs the impact of machine learning. Again, optimization has been around for longer and that certainly plays a role. And the impact of machine learning will continue to grow for sure. For those who don't believe you can point to many industries (airline industry being one) that have been built on the back and been absolutely transformed by optimization.
>
> **Answer (Mike):** What we are trying to do is trying to define this whole field of basically AI as a big umbrella term where optimization fits in there, machine learning fits in there and deep learning fits in there. All these algorithms fit in there and it's just a collection of algorithms people can use to automate and make better decisions with the data that they have. We have had some success in that and we have started to see some trends where the people who are heavy into Machine Learning and Data Science are discovering Linear Programming and what it can do to help them solve particular problems. In the end, no matter how you classify it all, optimization is an algorithm that should be in every data scientist's toolkit.

Question: When setting up optimization problems, what common constraints have you run up against that would have directly impacted success (or lack thereof)?

Answer (Ed): As a tool vendor, Gurobi provides the tools to the folks who are actually building the models. We get some exposure to how it's actually being used but this question is probably one Mike is closer to being able to give some good insights into.

Answer (Mike): We talked a lot about MIPs in this session. We encourage people to think about 'can we minimize the number of integer variables that you really need to have to get started?'. The initial reaction is to throw almost too much into the model. The second problem people run into is they don't think about what to do when there is an infeasible model. Some constraints, like minimization constraints, tend to come back with just infeasibility messages rather than point to where you are tight. We think it's important to have 'safety valves' when you are setting up your model so that it always runs. This is sometimes done with things we call slack variables but there are also a few other ways. The overarching idea is, 'how do you make it broad enough so that when I hit the run button I come up with some answer (that answer might not be feasible or practical) that at least tells me/points me in the direction of what's wrong with my model. I think that optimization around excitement will pick up once the ML community starts to learn more about it.

Question: How about stochastic optimization and its current state of practicality?

Answer (Ed): George Danzig back int he 1950s developed the Simplex method but from what I understand the problem he was really trying to solve was stochastic optimization. He worked on it for probably 30 years. From what we have seen from the application side is that it's very rare for an application to include a stochastic component. The explanation in almost all cases is 'ultimately you've got a budget for how much time you can spend on your optimization. So you have a choice... you can either add more detail to a deterministic model or you can add a stochastic component to it. So far almost everyone is deciding to add more detail.

Answer (Mike): Some of it is probably because the probabilities are hard to estimate. And even if you get the probabilities in for simple models its actually hard to interpret the results. What does it actually tell me? And do I believe it?

Question: What advice to you have for the students on this call looking to become optimization experts?

Answer (Ed): My main is advice is that the applications are the interesting stuff. It's useful to have a background in simplex method, branch and bound, cutting planes, that sort of thing. But you very quickly have what you need to be a user of the technology. To be an expert is to really be exposed to lots of different models that have been built. It's being able to walk into a situation and see the data that is available, the things that are driving profit, cost, etc (the things you are trying to optimize) and determine how you can build a model that will capture that without requiring a year or two to actually implement and deploy. Basically... expose yourself to lots of different optimization applications. The number of fundamentally different models that are out there is not that big. There are probably 10-20 different types of models. But there is a lot of cleverness in figuring out 'for this problem, for this customer, how am I going to take this model and modify it in this way in order to best solve the problem they are looking to tackle.'

Question: What about stochastic optimization for inventory control with demand forecasting uncertainty?

Answer (Mike): I think people are already doing this. Standard inventory optimization models already account for uncertainty around supply and demand. Even more exciting, we are seeing a trend for even more advanced inventory models that use techniques like reinforcement learning as part of the optimization. The Beer Game (https://beergame.opexanalytics.com/#/) gives you a glimpse into this.

Question: Any recommended papers/textbooks with interesting MIP model formulations?

Answer (Ed): We maintain a list of recommended resources here: https://www.gurobi.com/resource/books-blogs/

Question: Could you please elaborate on the speed and cores?

Answer (Ed): Over our broad MIP test set, we get right around a 2X average performance improvement from using 4 cores. I believe the number is just under 3X for 8 cores. MIP performance definitely doesn't scale linearly with cores if you look across a broad set of models.

Question: Given the explosion in data science in the business, what are some best practices, if you will, to leverage that as an optimization expert?

> **Answer (Ed):** Given how many more data scientists there are in industry than optimization experts, one big opportunity is just to make them more aware of optimization and help them to recognize optimization problems when they see them.
>
> **Answer (Mike):** We see that data science teams can enhance the value they bring to the business by having optimization as part of their toolkit. For example, after a great ML model gives you better predictions, you may need an optimization model to help you decide what to. The book "Prediction Machines" does a great job of saying the prediction models need judgement to make a decision. In our mind, optimization models can often fill the role of judgement.

Question: Have you seen folks confuse optimization with forecasting using ML?

> **Answer (Ed):** The main confusion we've seen is just a belief that ML can handle complex constraints. ML is very bad at solving linear systems, which is an essential building block for a lot of optimization models.

Question: As modelers, we end up having to add a lot of slack variables to equality constraints to diagnose infeasibilities. This may be okay from a performance standpoint and easier for constraints in continuous variables but gets challenging for IP constraints. While facilities like IIS are useful for developers and modelers, not so much for end-users and for deployment. Any thoughts on how to address this?

> **Answer (Ed):** Unfortunately, I think that soft constraints with penalty terms are the current best practice here.

Question: How do you manage general confusion around the word 'optimization' when someone refers to things like Database Optimization or Website Optimization?

> **Answer (Ed):** We spend a lot of time discussing this. At this point, we've pretty much given up on using 'optimization' alone. Mathematical optimization, or even just MIP, seem like better terms.

Question: Is Gurobi thinking about adding a constraint programming solver (such as CP Optimizer) any time soon?

Answer (Ed): No. We haven't seen a lot of demand, and we haven't had a lot of positive experiences with CP.

Question: How do users perceive benefits of using optimization solutions? What are some interesting ways that we can help users realize benefits of optimization solutions?

Answer (Mike): There are many ways to answer this question. I'll take a stab at just one way: you need to present the results in a way that business users and use the results. Often, this means you need a good interface to the solution that allows the user to interact with solutions.

Question: Machine Learning + Optimization has been on for some time now, but why don't we see many business applications?

Answer (Ed): We're actually seeing quite a few, although the integration is typically quite loose. In some cases, an existing MIP model may contain forecasted data, and the modeler may improve the forecast by adding an ML model.

Answer (Mike): We see a growing number of use cases. These can be some nice solutions where you make a prediction and use the optimization to help decide what to do.

Question: Metaheuristics are a very powerful tool for solving optimization problems, but maybe they are misused? Why do you think that is? Why the has MIP, LP etc become popular again vs the use of metaheuristics?

Answer (Ed): Meta-heuristics are definitely still quite popular. They have a few significant limitations that MIP can often address, including the lack of a quality guarantee and robustness issues when new constraints are added. I suspect that their popularity varies by community. I think Computer Scientists are less aware of the potential advantages of MIP, for example, so are more likely to use meta-heuristics.

Answer (Mike): Another important point is that MIP solvers continue to get faster every year (and a lot faster). So, what once needed a heuristic, the MIP can now solve.