Optimization and Analytics

Robert E. Bixby

Gurobi Optimization

The Early History

- 1947 – George Dantzig invents simplex for LP
  - Gives a global view of solutions
- 1951 – First computer code for solving LPs
- 1960 – LP commercially viable
- 1970 – MIP commercially viable

Gurobi Optimization

The Decade of the 70's

- **Interest in optimization flowered**
  - Numerous new applications identified
  - Many companies created OR Departments
- **Significant difficulties emerged**
  - Building application was very expensive and very risky
    - 3-4 year development cycles
    - Technology just wasn't ready: LPs were hard and MIP was a disaster
  - **Result:** *Disillusionment and much of that disillusionment persists to this day.*

Gurobi Optimization

The Decade of the 80's

- Several key developments
  - IBM PC introduced in 1981
  - Relational databases developed
    - ERP systems introduced
  - First algebraic modeling language: GAMS
  - Karmarkar's 1984 paper on interior-point methods

Gurobi Optimization

The Decade of the 90's

- LP performance takes off
- Data became plentiful and accessible
  - ERP systems became commonplace
- Optimization shown to be feasible on selected, difficult, real problems
  - Business: Airlines, Supply-Chain
  - Academic: Traveling Salesman Problem

Gurobi Optimization

Mixed Integer Programming

A Definition

A *mixed-integer program* (MIP) is an optimization problem of the form

$$\begin{aligned} &\textit{Minimize} && c^T x \\ &\textit{Subject to} && Ax = b \\ & && l \leq x \leq u \\ &\text{some or all } x_j \text{ integer} \end{aligned}$$

Gurobi Optimization

1998 … A New Generation of MIP Codes

- Linear programming
  - Stable, robust dual simplex
- Variable/node selection
  - Influenced by traveling salesman problem
- Primal heuristics
  - 12 different tried at root
  - Retried based upon success
- Node presolve
  - Fast, incremental bound strengthening (very similar to Constraint Programming)
- Presolve – numerous small ideas
  - Probing in constraints:

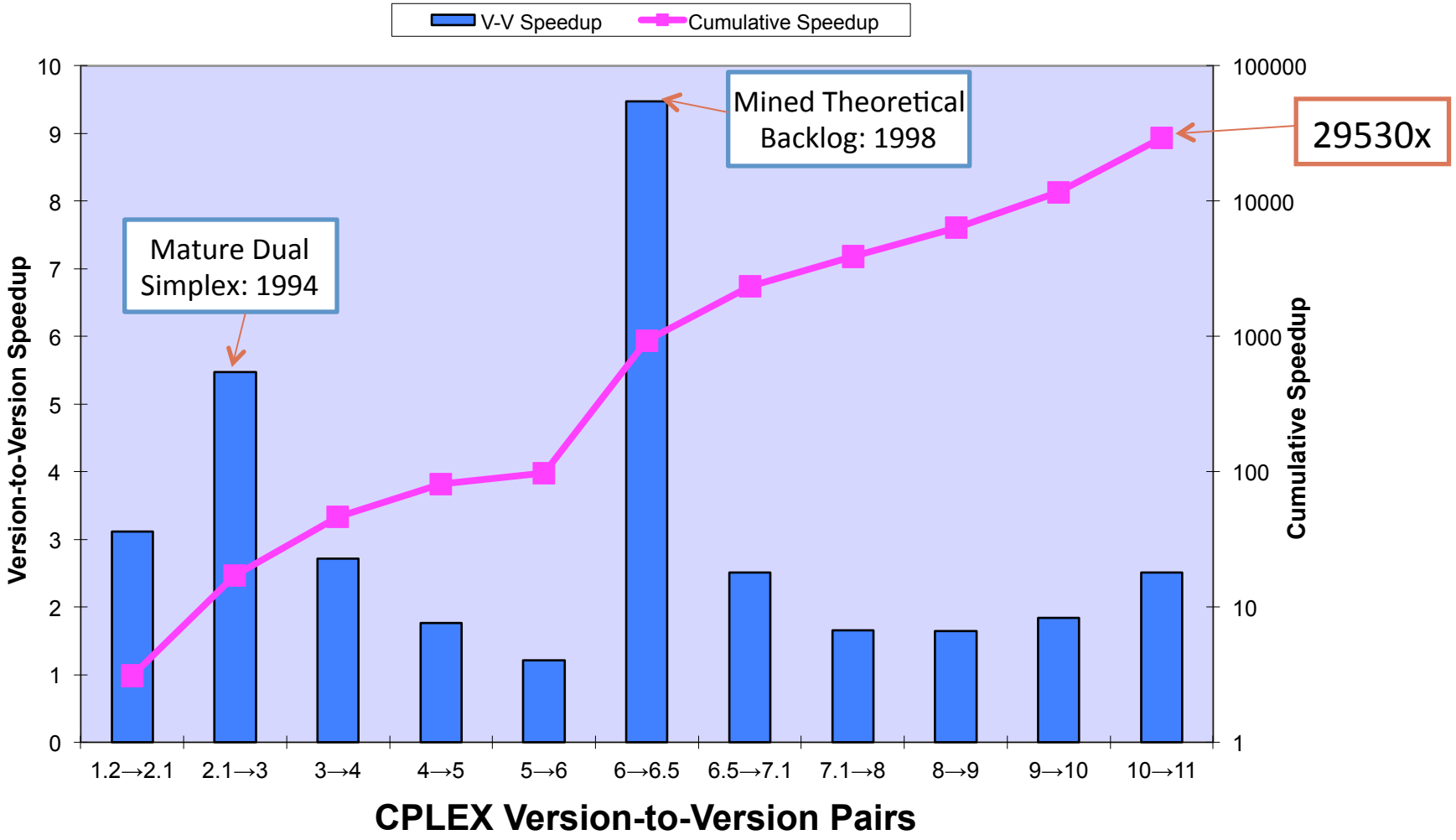
    $\sum x_j \leq (\sum u_j)\, y, \quad y = 0/1$

    ➔ $x_j \leq u_j y$ (for all j)
- Cutting planes
  - Gomory, mixed-integer rounding (MIR), knapsack covers, flow covers, cliques, GUB covers, implied bounds, zero-half cuts, path cuts

Gurobi Optimization

MIP Performance Improvements

Speedups 1991-2008



Progress: 2009 – Present

MIP Speedup 2009-Present

- Starting point
  - Gurobi 1.0 & CPLEX 11.0 ~equivalent on 4-core machine

- Gurobi Version-to-version improvements
  - Gurobi 1.0 -> 2.0: 2.4X
  - Gurobi 2.0 -> 3.0: 2.2X (5.1X)
  - Gurobi 3.0 -> 4.0: 1.3X (6.6X)
  - Gurobi 4.0 -> 5.0: 2.0X (12.8X)
  - Gurobi 5.0 -> 6.0: 2.2X (27.6X)

- Machine-independent IMPROVEMENT since 1991
  - Over 800,000X -- 1.8X/year

Gurobi Optimization

Big Data: Million-Sized Models

- **One view:** Optimization is a technique for analyzing/understanding large data sets and producing actionable conclusions.
- Customer regularly send us models
  - Over 300 customers have sent models
  - More than 10,000 models in total
- Trend: Models are getting larger and larger
  - One of our customers has recently "reported" a model with 2 billion nonzeros.

Gurobi Optimization

MIP "million sized" models

- Customer models with > 1 million variables or constraints
  - 2011: 47 models
  - 2012: 68 models
  - 2013: 156 models
  - 2014: 286 models
  - From energy, transportation, finance, mining, forestry, social media, …
- Largest
  - Trading – stock market
    - 39M rows, 39M columns, 100M nonzeros
    - Solution time: 1040 seconds
- Testing (2014 – 156 models)
  - Ran defaults
  - 7200 second time limit

Gurobi Optimization

Summary – 156 Models in Total

- **Solvability – defaults (tuned): 144**
  - Solved to optimality 81 (88)
  - Feasible < 10% gap 27 (33)
  - Feasible > 10% gap 12 (23)
- Others – no feasible solution: 12
  - Could not solve root LP: 2
- **Summary**
  - Defaults: 70% acceptable solutions
  - Tuned: 78% acceptable solutions
  - No feasible solution found: 8% found
  - Mean solve time for solvable: 496 sec.

Gurobi Optimization

The Curse of the Free Solver

Mittelmann MIPLIB2010 tests (87 models):

	Factor Slower than Gurobi	% Solved
GLPK	27X	1%
LPSOLVE	24X	6%

The Curse: These solvers ***very popular*** (in particular in statistics) and lead to ***completely wrong conclusions*** about what is doable!

Gurobi Optimization

Statistics and Prescriptive Analytics

- Dan Gusfield
  - Statistical applications in computational biology
  - web.cs.ucdavis.edu/~gusfield
- Dimitris Bertsimis
  - Application of MIP to regression methods
  - "Best subset selection via a modern optimization Lens", with A. King and R. Mazumder
- Stan Uryasev – Finance
  - Statistical applications in finance
  - www.ise.ufl.edu/uryasev/research/testproblems/advanced-statistics
- José R. Zubizarreta
  - MISMATCH: **optimal matching in observational studies using MIP**
  - www.columbia.edu/~jz2313

Gurobi Optimization

The Analytics Wave

The Analytics Taxonomy

(Thomas Davenport)

- *Descriptive Analytics*
  - Often what people mean when they say "Big Data"
  - Understanding, representing your data
- *Predictive Analytics*
  - Forecasting
  - Making predictions based on your data
  - Statistics plays a central role
- *Prescriptive Analytics*
  - Using your data to make decisions
  - Optimization plays a central role
    - Mixed-Integer Programming (MIP) is THE primary optimization technique used in business.

Gurobi Optimization

Remarks I: The Future of Analytics and Optimization

- We should ride the wave
- Analytics is much bigger than optimization
  - Analytics has real chance of becoming a "standard" business function
  - Concretely: Companies of the future may well have *Chief Analytics Officers*
- Our objective should be: Profit from this movement
  - Make sure we are part of the analytics education process
  - View corporate analytics groups as providing a natural home for the OR / Management Science function

Gurobi Optimization

Remarks II: The Importance of Optimization

- Making the right DECISIONS based on your data is the first priority in managing a modern enterprise
  - Optimization is about making decisions!
- Many people say prescriptive analytics follow predictive analytics.
  - In my view this puts the cart before the horse: There is no better tool than building a model to force you to understand your data and determine what data you need

Thomas H. Davenport
"Analytics 3.0", HBR, December 2013

"Although Analytics 3.0 includes all three types [descriptive, predictive, prescriptive], it emphasizes the last. Prescriptive models involve large-scale testing and optimization and are a means of embedding analytics into key processes and employee behaviors."

Gurobi Optimization