

Modeling and Solving Discrete Optimization Problems in Practice

Juan Pablo Vielma

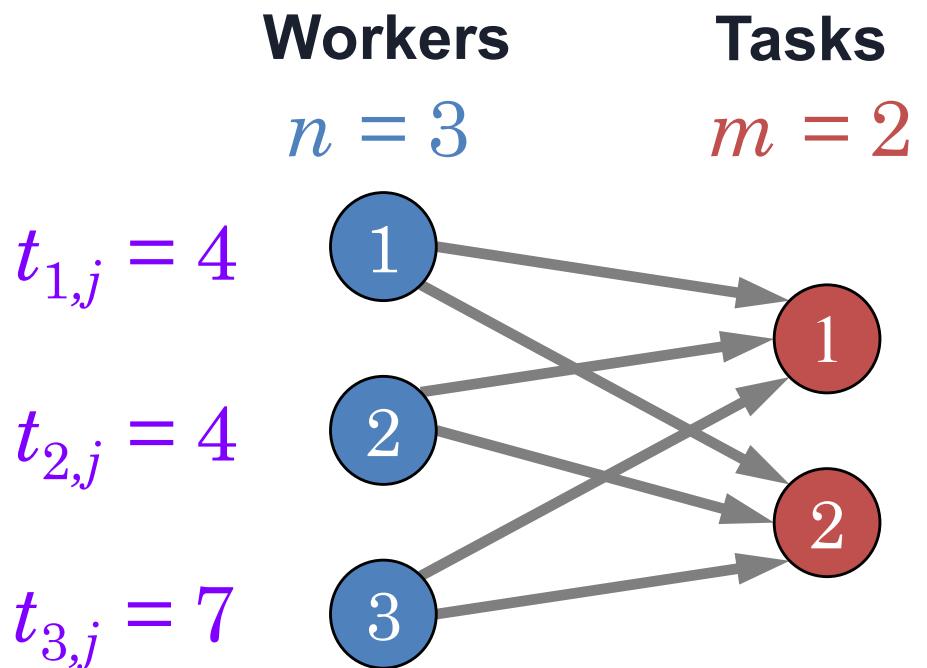
Massachusetts Institute of Technology

18.095 - Mathematics Lecture Series.
Cambridge, MA, IAP 2018.

Combinatorial Example: Assignment Problem

- Assign n workers to m tasks to complete all tasks
- At most one task per worker
- Worker i takes $t_{i,j}$ hours to complete task j
- Minimize total time worked

- Graph:
 - Worker and task **nodes**
 - Arcs between worker and task **nodes**



Combinatorial Example: Assignment Problem

- Assign n workers to m tasks to complete all tasks
- At most one task per worker
- Worker i takes $t_{i,j}$ hours to complete task j
- Minimize total time worked
- Variables: $x_{i,j} = 1$ if worker i is assigned to task j and 0 o.w.

$$\min \quad \sum_{i=1}^n \sum_{j=1}^m t_{i,j} x_{i,j}$$

s.t.

$$\sum_{j=1}^m x_{i,j} \leq 1 \quad \forall i \in \{1, \dots, n\} \quad \text{Worker constraints}$$

$$\sum_{i=1}^n x_{i,j} \geq 1 \quad \forall j \in \{1, \dots, m\} \quad \text{Task constraints}$$

$$x_{i,j} \in \{0, 1\} \quad \forall i \in \{1, \dots, n\}, j \in \{1, \dots, m\}$$

Traveling Salesman Problem : Visit all Cities Once

Firefox File Edit View History Bookmarks Tools Window Help Sun 1:12 PM Google Maps http://maps.google.com/ Gmail Google Notebook La Tercera Apple Insider Currency Converter Web Images Maps News Shopping Gmail more mgoyc00l@gmail.com | My Profile | Saved Locations | Help | Web History | My Account | Sign out

Start address e.g. "SFO" End address e.g. "94526" Get Directions

Search Results My Maps

8,970 mi – about 5 days 22 hours

A From: I-5 N Show all directions

B To: US-310

C To: I-94 E

D To: US-36

E To: I-72 E

F To: I-94 W/US-41 N

G To: I-69 N

H To: RT-9

From: I-82 W @45.808880, -119.383310 To: US-310 @44.913820, -108.611000 t

Get Directions Street View Traffic Map Satellite Terrain

Print Send Link to this page

Map showing a route from I-82 W (@45.808880, -119.383310) to US-310 (@44.913820, -108.611000). The route is highlighted in blue and passes through several states and provinces, including California, Oregon, Washington, Idaho, Montana, North Dakota, Minnesota, Iowa, Missouri, Kansas, Nebraska, Oklahoma, Texas, Louisiana, Mississippi, Alabama, Georgia, South Carolina, North Carolina, Virginia, Maryland, Delaware, New Jersey, Connecticut, Rhode Island, Massachusetts, New Hampshire, Vermont, New York, and Canada. The total distance is 8,970 mi and it takes about 5 days 22 hours.

500 mi 500 km North Pacific

©2008 Google - Map data ©2008 LeadDog Consulting, NAVTEQ™, Europa Technologies - Terms of Use

Formulation for Traveling Salesman Problem

$$[n] := \{1, \dots, n\}$$

$$\min \sum_{i,j=1}^n d(i,j)x_{i,j}$$

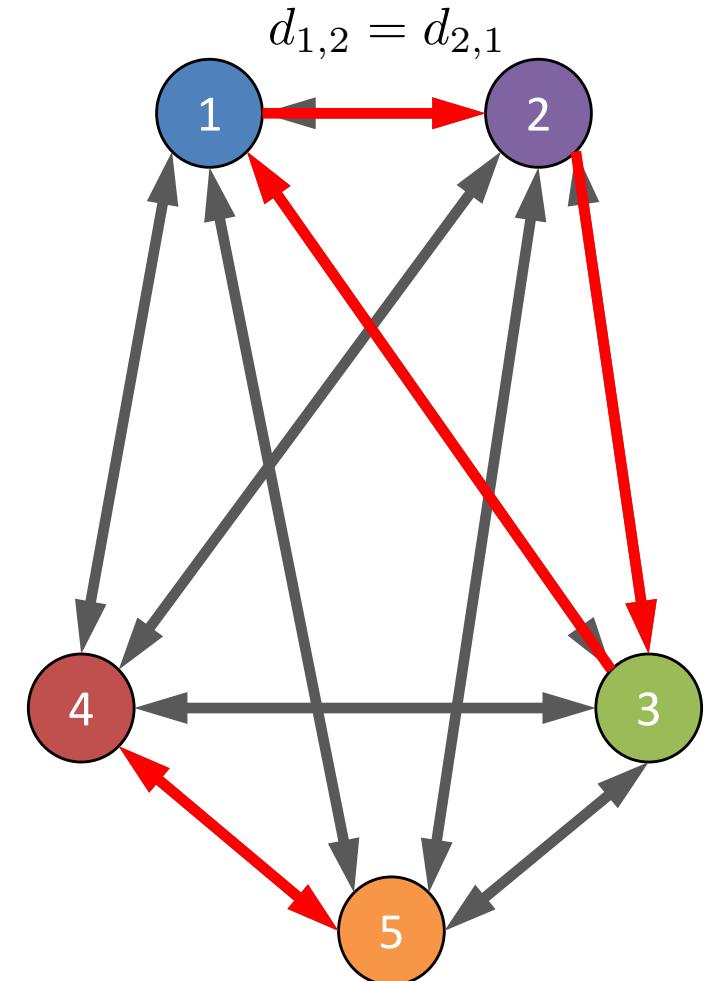
s.t.

$$\sum_{j=1}^n x_{i,j} = 1 \quad \forall i \in [n]$$

$$\sum_{i=1}^n x_{i,j} = 1 \quad \forall j \in [n]$$

$$x_{i,i} = 0 \quad \forall i \in [n]$$

$$x_{i,j} \in \{0, 1\} \quad \forall i, j \in [n]$$

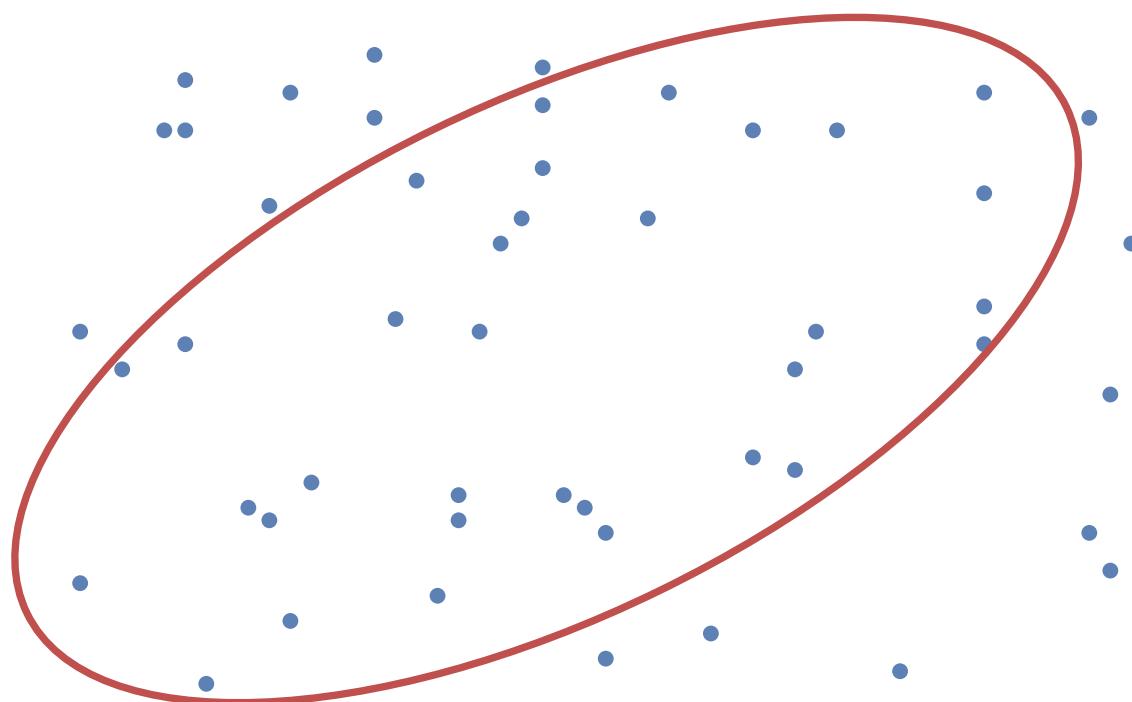


Homework Question 1: Add missing constraints

Hint: You will need around 2^n inequalities

Mixed Integer Programming (MIP)

- Discrete and continuous variables or combinatorial constraints on continuous variables.
- Example: Find minimum volume ellipsoid that contains 90% of data points



MIP & Daily Fantasy Sports



> \$15K

Download Code from Github:

<https://github.com/dscotthunter/Fantasy-Hockey-IP-Code>

<http://arxiv.org/pdf/1604.01455v1.pdf>

How hard is MIP: Traveling Salesman Problem ?

The screenshot shows a Firefox browser window with the following details:

- Title Bar:** Firefox, File, Edit, View, History, Bookmarks, Tools, Window, Help.
- Address Bar:** http://maps.google.com/
- Google Maps Search Results:** The search term "Paradoxes, Contradictions, and the Limits of Science" was entered. The results page is visible on the right.
- Left Panel (American Scientist Magazine Cover):**
 - Header:** AMERICAN Scientist, May–June 2016, www.americanscientist.org
 - Section Headings:** How scientists can avert a NEW FLINT CRISIS, The urgent need for DROUGHT-PROOF ENERGY, Exploring the landscape of COSMIC HABITATS.
 - Main Article:** Cyber-Insecurity, The latest digital threats call for a smarter, stronger response.
 - Sigma Xi Logo:** THE SCIENTIFIC RESEARCH SOCIETY

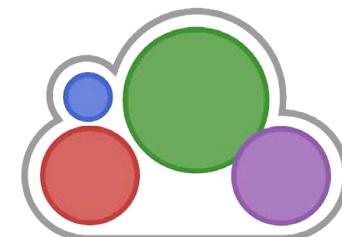
MIP = Avoid (Complete) Enumeration

- Number of tours for 49 cities = $48!/2 \approx 10^{60}$
- Fastest supercomputer $\approx 10^{17}$ flops
- Assuming one floating point operation per tour:
 $> 10^{35}$ years $\approx 10^{25}$ times the age of the universe!
- How long does it take on an iphone?
 - < 1 sec ! Dantzig, Fulkerson and Johnson 🖊 in 54'
 - Even theoretically hard MIPs “can” be solved:
 - Open-source solvers: GLPK, CBC, etc.
 - Commercial: Gurobi, CPLEX, etc.
 - Modeling Language:



Easy MIP through **julia** & **JuMP**

- **julia** : general purpose programming language
 - download <https://julialang.org/downloads/>
then click or run from command line
- **JuMP** : modeling language for optimization
-  **GLPK** : Open-source MIP solver
 - `julia> Pkg.add("JuMP"); Pkg.add("GLPKMathProgInterface")`
- Can also try JuliaBox on web
 - <https://www.juliabox.com/>



- Assignment problem:

$$\min \quad \sum_{i=1}^n \sum_{j=1}^m t_{i,j} x_{i,j}$$

s.t.

$$\sum_{j=1}^m x_{i,j} \leq 1 \quad \forall i \in \{1, \dots, n\}$$

$$\sum_{i=1}^n x_{i,j} \geq 1 \quad \forall j \in \{1, \dots, m\}$$

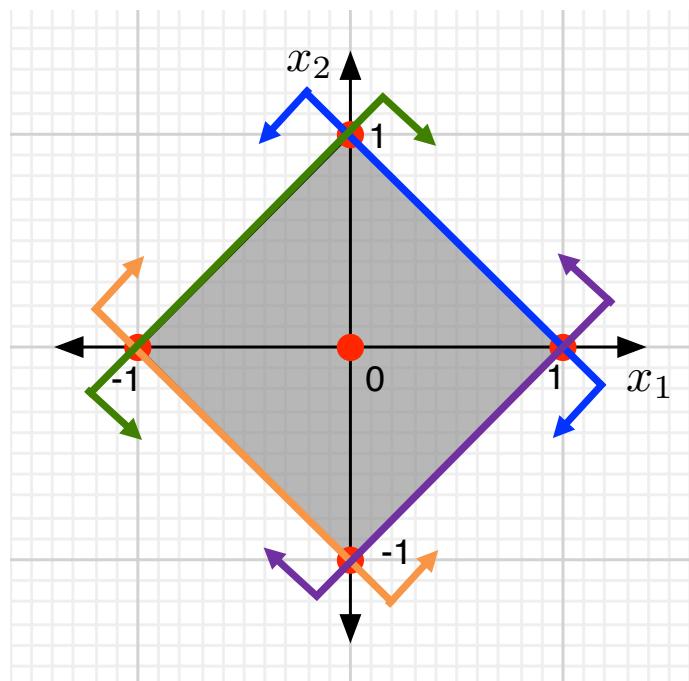
$$x_{i,j} \in \{0, 1\} \quad \forall i \in \{1, \dots, n\}, j \in \{1, \dots, m\}$$

```
model = Model(solver=GLPKSolverMIP());
@variable(model, x[1:n,1:m], Bin);
@objective(model, Min, sum(t[i,j]*x[i,j] for i in 1:n, j in 1:m));
@constraint(model, [i=1:n], sum(x[i,j] for j in 1:m) <= 1);
@constraint(model, [j=1:m], sum(x[i,j] for i in 1:n) >= 1);
```

Homework Question 2: Solve problem with random cost
 Complete file in website.

Solving MIPs: Step 1 = Linear Programming

$$\begin{aligned} \max \quad & x_2 \\ \text{s.t.} \quad & x_1 + x_2 \leq 1 \\ & -x_1 - x_2 \leq 1 \\ & +x_1 - x_2 \leq 1 \\ & -x_1 + x_2 \leq 1 \\ & x_1, x_2 \in \mathbb{Z} \end{aligned}$$



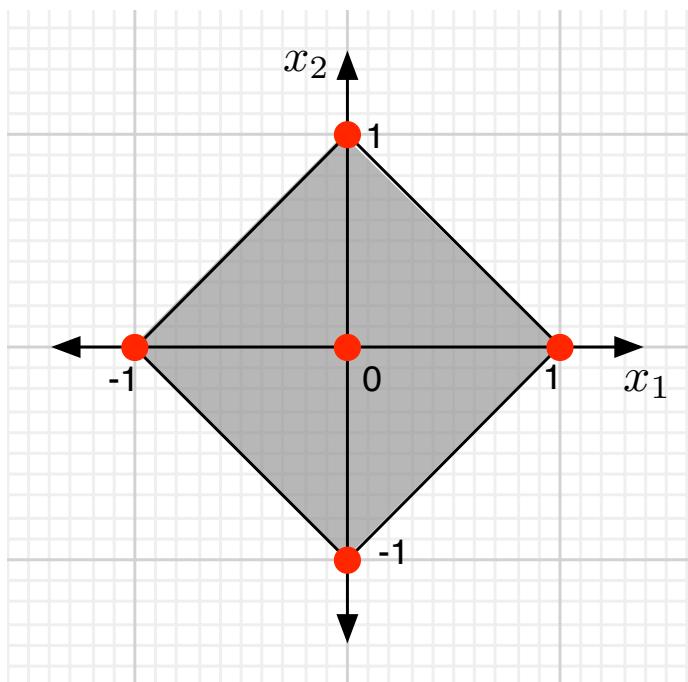
← Linear Programming (LP) Relaxation

- Solving LPs is easy in theory and practice.
- One reason = LP duality
 - Suppose I guess optimum $x_1 = 0$ and $x_2 = 1$.
 - How do I prove that for all solutions of LP $x_2 \leq 1$?

$$\begin{aligned} & (1/2) \times (x_1 + x_2 \leq 1) \\ & + (1/2) \times (-x_1 + x_2 \leq 1) \\ \hline & x_2 \leq 1 \end{aligned}$$

Solving MIPs: Step 1 = Linear Programming

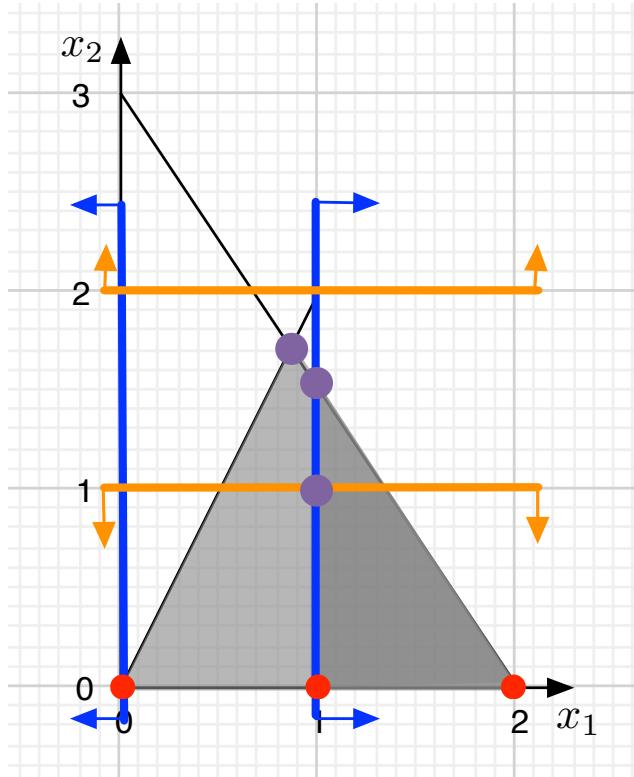
$$\begin{aligned} \max \quad & x_2 \\ \text{s.t.} \quad & x_1 + x_2 \leq 1 \\ & -x_1 - x_2 \leq 1 \\ & +x_1 - x_2 \leq 1 \\ & -x_1 + x_2 \leq 1 \\ & x_1, x_2 \in \mathbb{Z} \end{aligned}$$



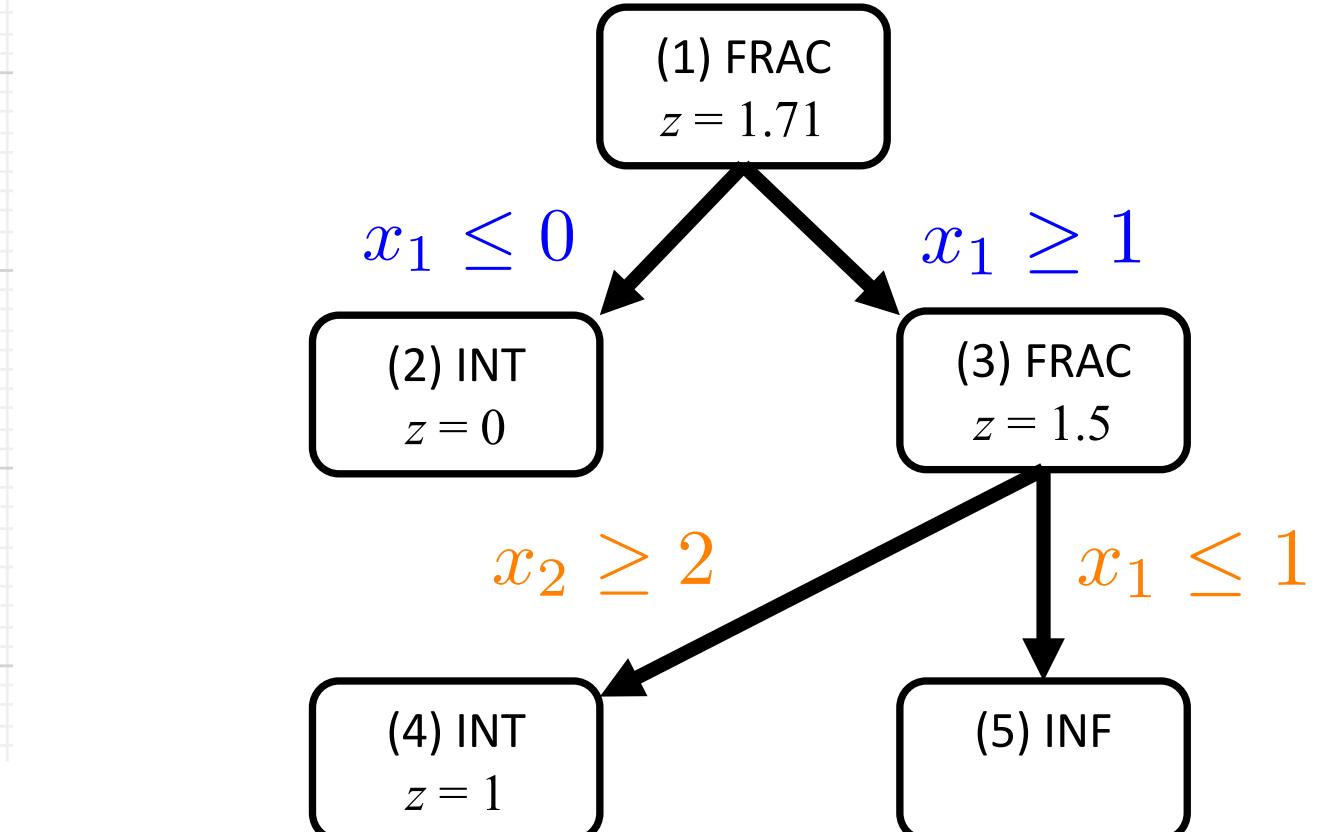
← Linear Programming (LP) Relaxation

- LP relaxation always gives a (upper) bound on the MIP:
 - If solution of LP is “integer” then you solved the MIP
 - LP solvers return “corner” solution, which fixes “multiple optima” (e.g. $\max x_1 + x_2$)
 - Homework Question 3: Solve LP relaxation of assignment problem with JuMP. Is solution integer?

Solving MIPs: Step 2 = Branch-and-Bound

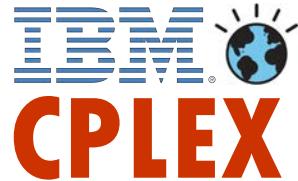


$$\begin{aligned} \max z &:= x_2 \\ 3x_1 + 2x_2 &\leq 6 \\ -2x_1 + x_2 &\leq 0 \\ x_1, x_2 &\geq 0 \\ x_1, x_2 &\in \mathbb{Z} \end{aligned}$$



← Linear Programming (LP) Relaxation
Homework Question 4:
Prove $x_2 \leq 12 / 7$ for LP Relaxation.

Modern MIP Solvers = B&B++



- Really branch-and-cut:
 - Use cuts to improve LP relaxation.
- Elaborate heuristics: Rounding ++
- Preprocessing: fixing variables by logical implications.
- Advanced management of B&B tree.
- Extensive tuning of parameters and techniques.

Cutting Plane Example: Chátal-Gomory Cuts

$$P := \left\{ x \in \mathbb{R}^2 : \begin{array}{l} x_1 + x_2 \leq 3, \\ 5x_1 - 3x_2 \leq 3 \end{array} \right\}$$

\cap

$$H := \left\{ x \in \mathbb{R}^2 : \underbrace{4x_1 + 3x_2}_{\in \mathbb{Z}} \leq 10.5 \right\}$$

\downarrow

if $x \in \mathbb{Z}^2$

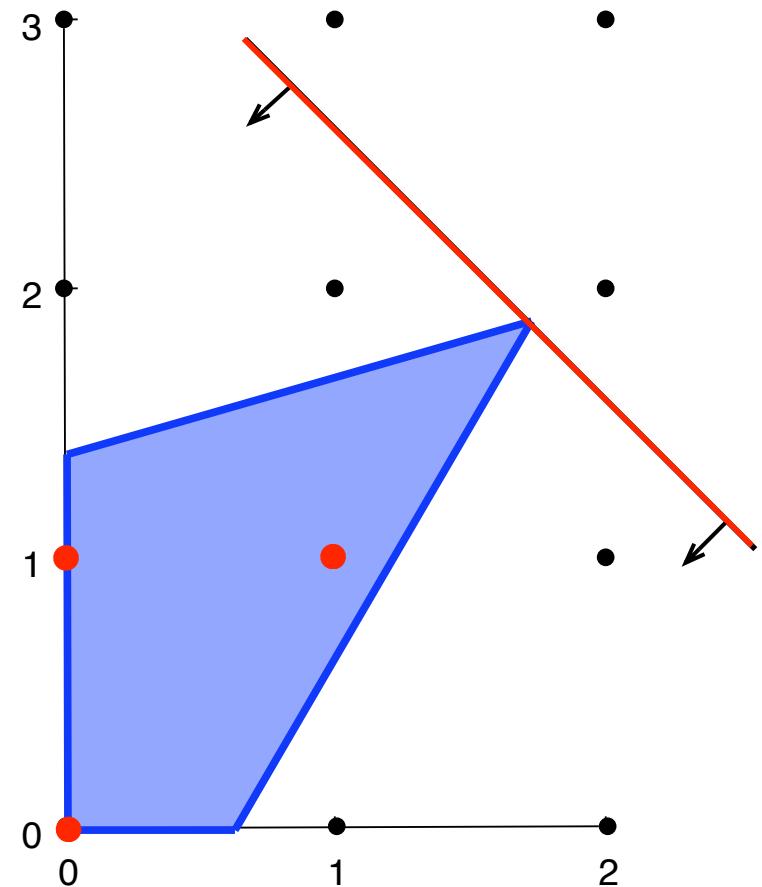
$$4x_1 + 3x_2 \leq [10.5]$$

Valid for $H \cap \mathbb{Z}^2$

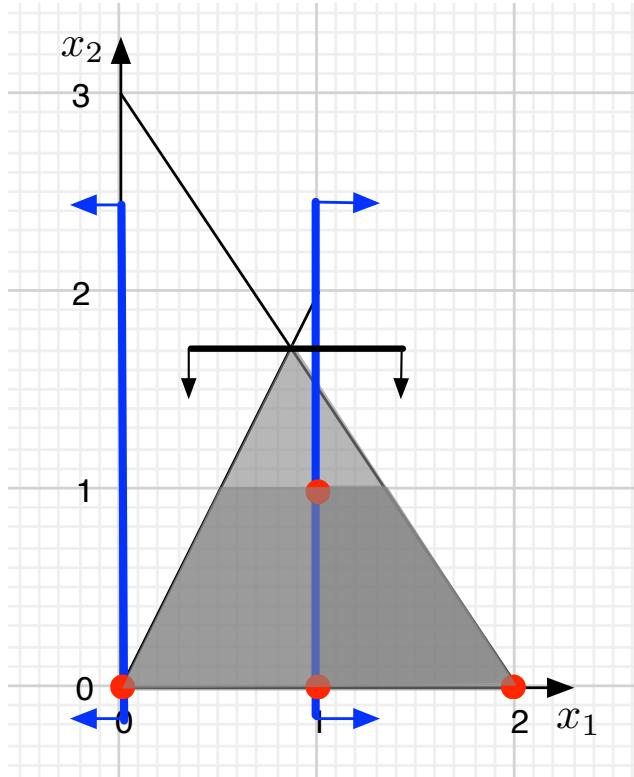
Valid for $P \cap \mathbb{Z}^2$

$$\begin{aligned} & (27/8)(x_1 + x_2 \leq 3) \\ + & (1/8)(5x_1 - 3x_2 \leq 3) \end{aligned}$$

$$\Rightarrow 4x_1 + 3x_2 \leq 10.5$$

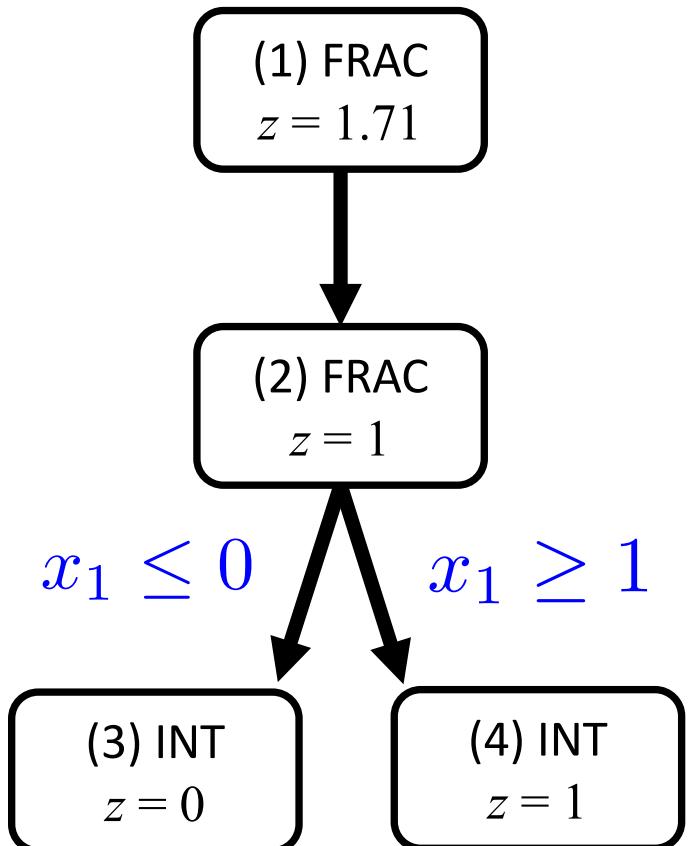


Branch-and-Bound and Cuts (Branch-and-Cut)

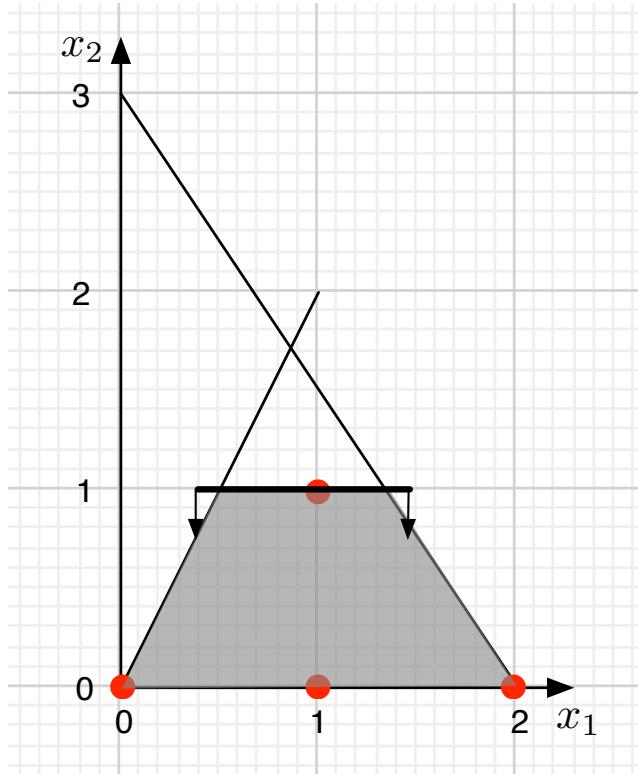


$$\begin{aligned} \max z &:= x_2 \\ 3x_1 + 2x_2 &\leq 6 \quad x_2 \leq \lfloor 1.71 \rfloor = 1 \\ -2x_1 + x_2 &\leq 0 \\ x_1, x_2 &\geq 0 \end{aligned}$$

$$x_1, x_2 \in \mathbb{Z}$$



Branch-and-Bound and Cuts (Branch-and-Cut)



$$\begin{aligned} \max z &:= x_2 \\ 3x_1 + 2x_2 &\leq 6 \quad x_2 \leq \lfloor 1.71 \rfloor = 1 \\ -2x_1 + x_2 &\leq 0 \\ x_1, x_2 &\geq 0 \end{aligned}$$

$x_1, x_2 \in \mathbb{Z}$

Homework Question 5:
Add two more Chátal-Gomory cuts so
the LP relaxation with all cuts solves
the MIP.

Hint:

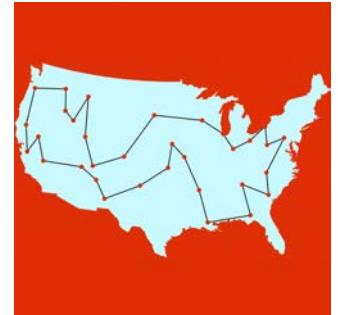
$$\begin{array}{r} (1/3)(3x_1 + 2x_2 \leq 6) \\ + (1/3)(\quad \quad \quad x_2 \leq 1) \\ \hline \end{array}$$

and

$$\begin{array}{r} (1/2)(-2x_1 + x_2 \leq 0) \\ + (1/2)(\quad \quad \quad x_2 \leq 1) \\ \hline \end{array}$$

No Enumeration = Keep Adding Cuts

- Number of tours for 49 cities = $48!/2 \approx 10^{60}$
- Fastest supercomputer $\approx 10^{17}$ flops
- Assuming one floating point operation per tour:
 $> 10^{35}$ years $\approx 10^{25}$ times the age of the universe!
- How long does it take on an iphone?
 - < 1 sec ! Dantzig, Fulkerson and Johnson 🖊 in 54'
 - This is how DFJ solved the problem by hand in 54'
 - In practice Branch-and-Cut is better.
 - More details in Concord TSP App
 - Cutting plane tutorial for TSP
 - <http://www.math.uwaterloo.ca/tsp/iphone/>



Easy Problems : LP Relaxation Always Integral

Consequence of LP duality: Kőnig's theorem

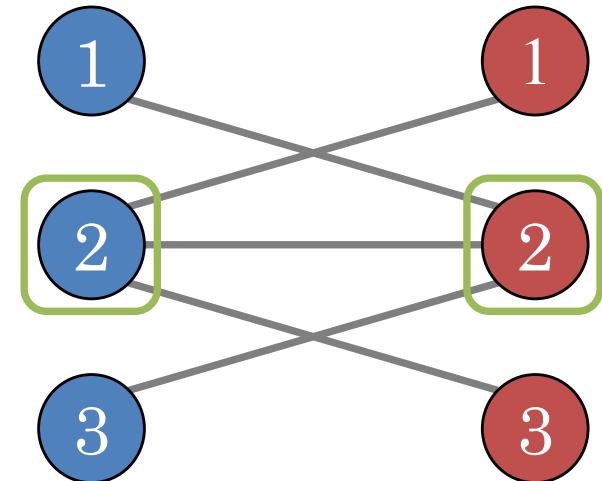
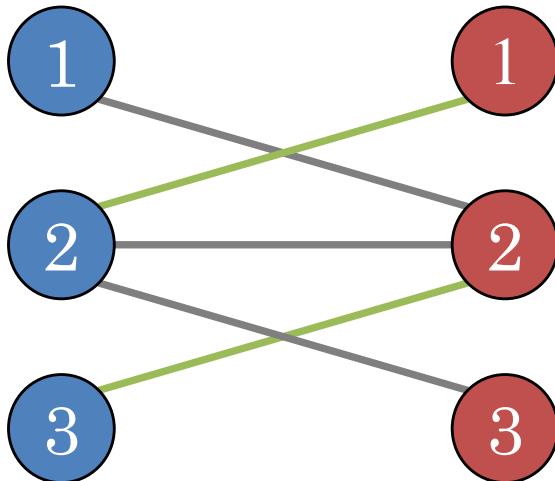
- Largest Matching

- Pick edges, at most one edge per node

=

- Smallest Node Cover

- Pick nodes that touch all edges



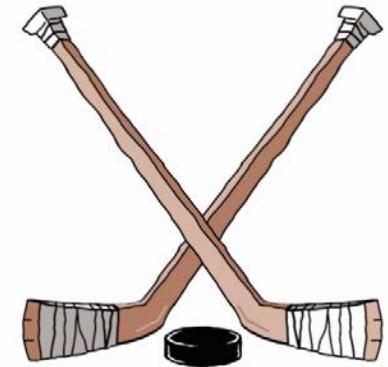
Classes and Links

- **julia** ,  **JuMP** and Optimization
 - <https://github.com/JuliaOpt/JuMP.jl>
 - <http://www.juliaopt.org>
- 15.053 Optimization Methods in Business Analytics
 - Modeling and computation
 - Instructor: James B. Orlin
 - Spring 2018: <http://mit.edu/15.053/www/>
- 18.453 Combinatorial Optimization
 - Theory and algorithms
 - Instructor: Michel Goemans
 - Spring **2017** : <http://www-math.mit.edu/~goemans/18453S17/18453.html>

MIP & Daily Fantasy Sports

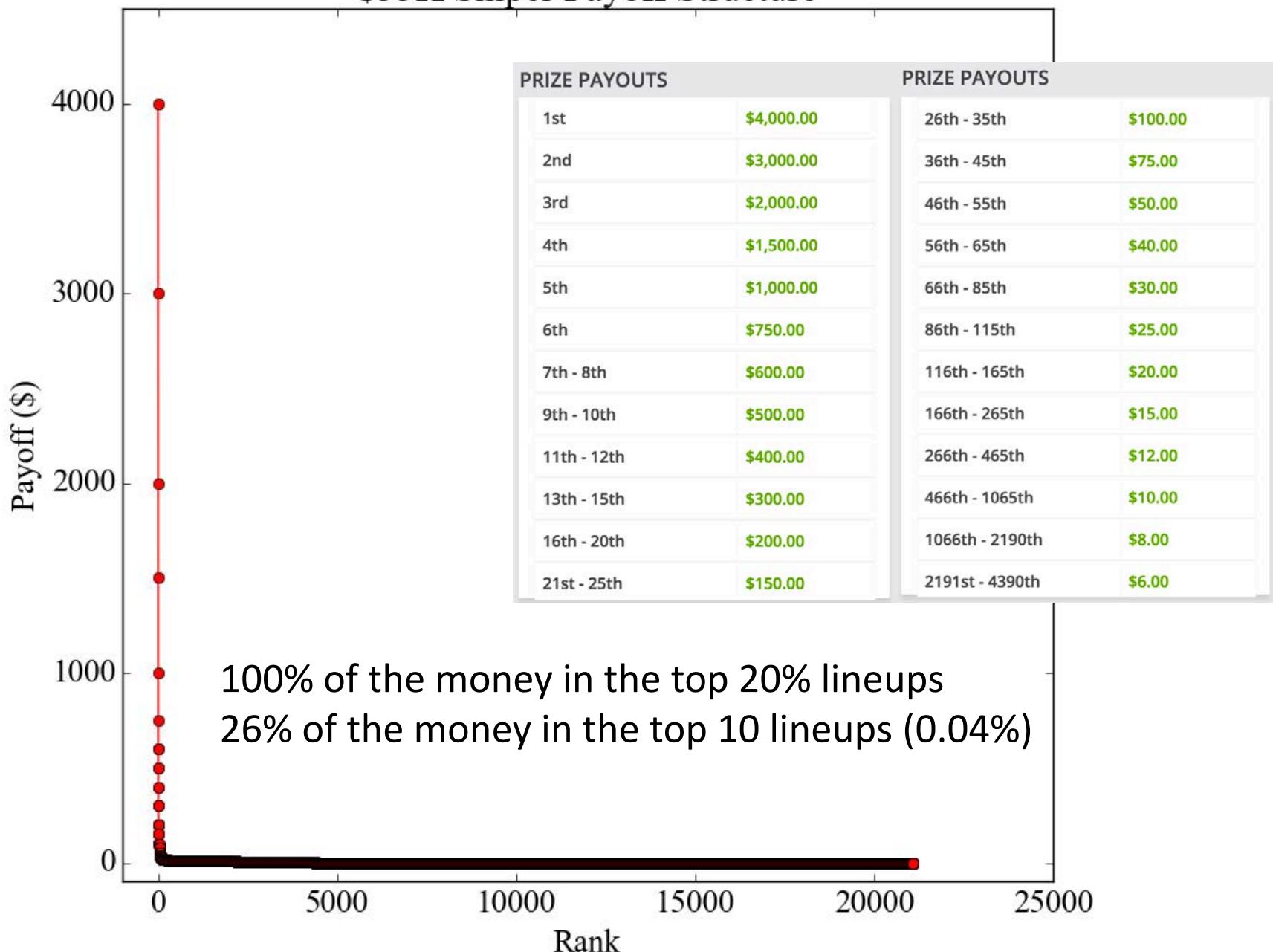


Example Entry



LINEUP					Avg. Rem. / Player: \$0	Rem. Salary: \$0
POS	PLAYER	OPP	FPPG	SALARY		
C	Jussi Jokinen	Fla@Anh	3.1	\$5,300	X	
C	Brandon Sutter	Pit@Van	3.0	\$4,400	X	
W	Nikolaj Ehlers	Wpg@Tor	3.9	\$4,800	X	
W	Daniel Sedin	Pit@Van	3.8	\$6,400	X	
W	Radim Vrbata	Pit@Van	3.4	\$5,800	X	
D	Brian Campbell	Fla@Anh	2.6	\$4,100	X	
D	Morgan Rielly	Wpg@Tor	3.5	\$4,200	X	
G	Corey Crawford P	StL@Chi	6.3	\$7,800	X	
UTIL	Blake Wheeler	Wpg@Tor	4.8	\$7,200	X	

\$55K Sniper Payoff Structure



Building a Lineup



MIP Formulation

- L lineups : indexed by l
- 9 players per lineup: indexed by p
- Decision variables

$$x_{pl} = \begin{cases} 1, & \text{if player } p \text{ in lineup } l \\ 0, & \text{otherwise} \end{cases}$$

Basic Feasibility

- Basic constraints:
 - 9 different players
 - Salary less than \$50,000

LINEUP		vg. Rem. / Player: \$0	Rem. Salary: \$0
POS	PLAYER	SALARY	
C	Jussi Jokinen	\$5,300	X
C	Brandon Sutter	\$4,400	X

 c_p

$$\sum_{p=1}^N c_p x_{pl} \leq \$50,000, \quad (\text{budget constraint})$$

$$\sum_{p=1}^N x_{pl} = 9, \quad (\text{lineup size constraint})$$

$$x_{pl} \in \{0, 1\}, \quad 1 \leq p \leq N.$$

Position Feasibility

- Between 2 and 3 centers
- Between 3 and 4 wingers
- Between 2 and 3 defensemen
- 1 goalie

Position constraints

$$2 \leq \sum_{p \in C} x_{pl} \leq 3, , \quad (\text{center constraint})$$

$$3 \leq \sum_{u \in W} x_{pl} \leq 4, \quad (\text{winger constraint})$$

$$2 \leq \sum_{u \in D} x_{pl} \leq 3, \quad (\text{defensemen constraint})$$

$$\sum_{u \in G} x_{pl} = 1 \quad (\text{goalie constraint})$$

Team Feasibility

- At least 3 different NHL teams

Team constraints

$$t_i \leq \sum_{p \in T_i} x_{pl}, \quad \forall i \in \{1, \dots, N_T\}$$

$$\sum_{i=1}^{N_T} t_i \geq 3,$$

$$t_i \in \{0, 1\}, \quad \forall i \in \{1, \dots, N_T\}.$$

Maximize Points

- Forecasted points for player p: f_p



Score type	Points
Goal	3
Assist	2
Shot on Goal	0.5
Blocked Shot	0.5
Short Handed Point Bonus (Goal/Assist)	1
Shootout Goal	0.2
Hat Trick Bonus	1.5
Win (goalie only)	3
Save (goalie only)	0.2
Goal allowed (goalie only)	-1
Shutout Bonus (goalie only)	2

Table 1 Points system for NHL contests in DraftKings.

Points Objective Function

$$\sum_{p=1}^N f_p x_{pl}$$

Lineup

Projections: 5.4 2.5 3.4 3.0 3.2 4.2 3.5 3.4 5.7

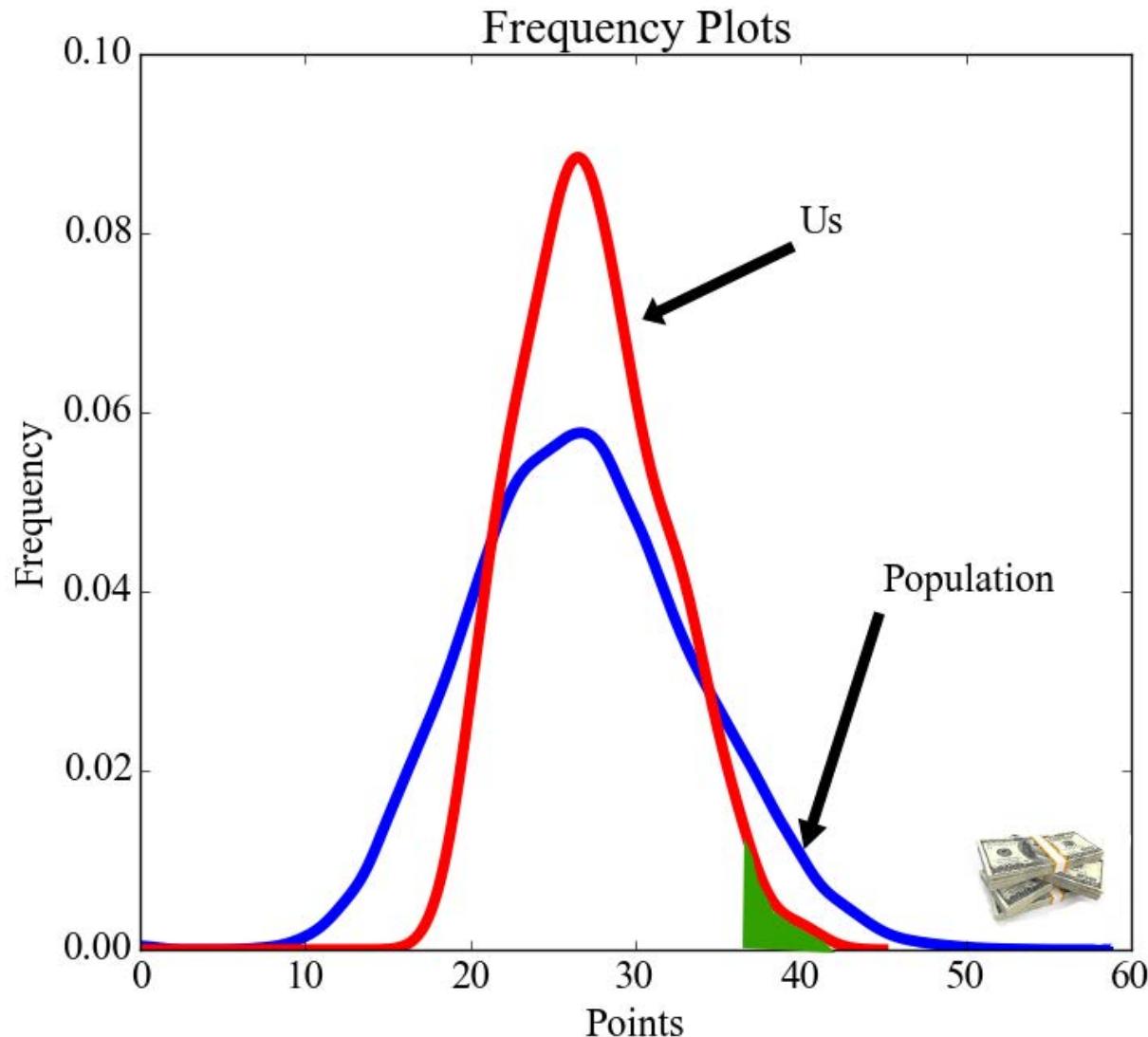
\$9500 \$2700 \$4600 \$3800 \$4600 \$6400 \$5200 \$5100 \$8000

W UTIL D D C C W W G

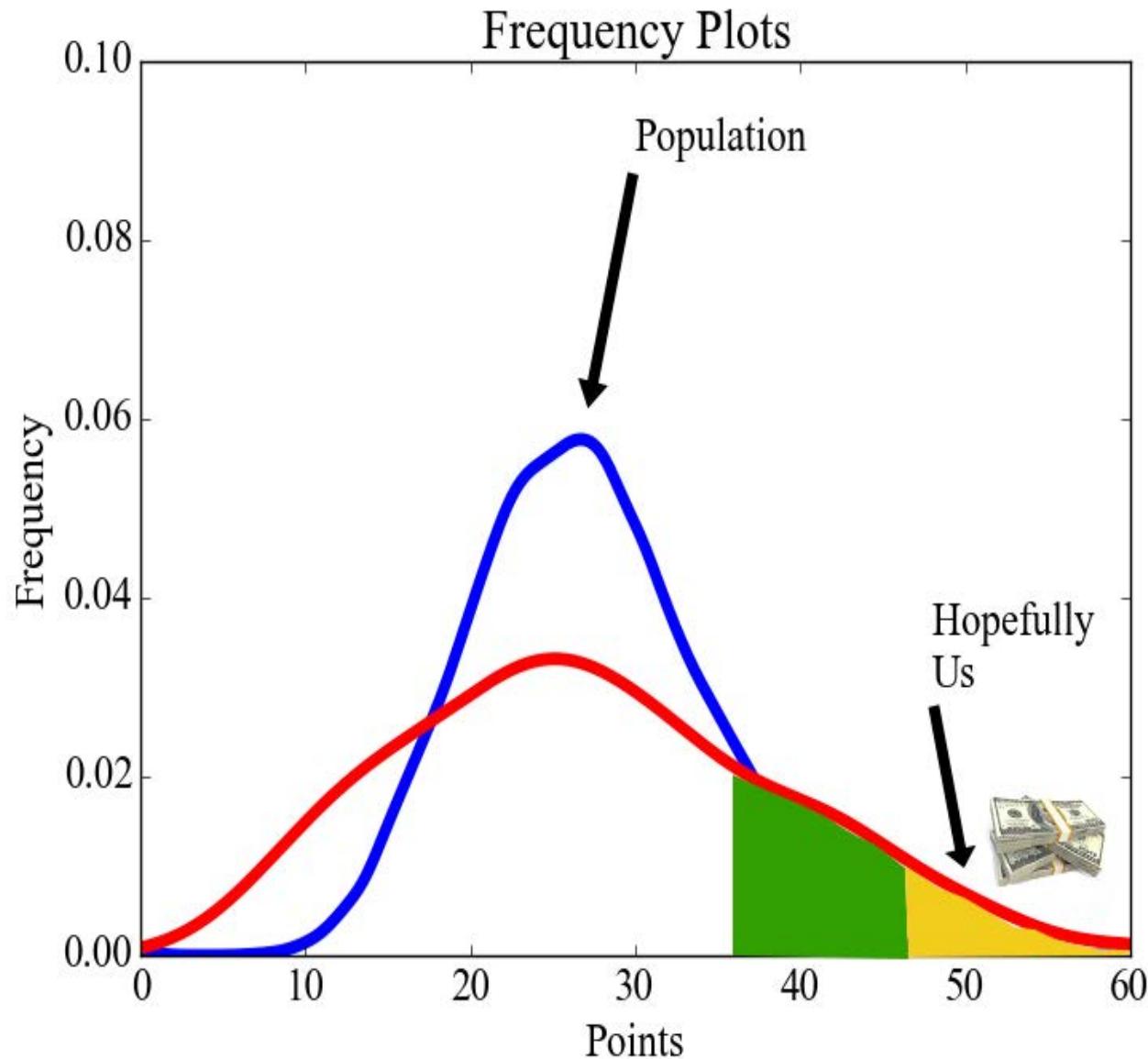


23 points on average

Need > 38 points for a chance to win



Increase variance to have a chance

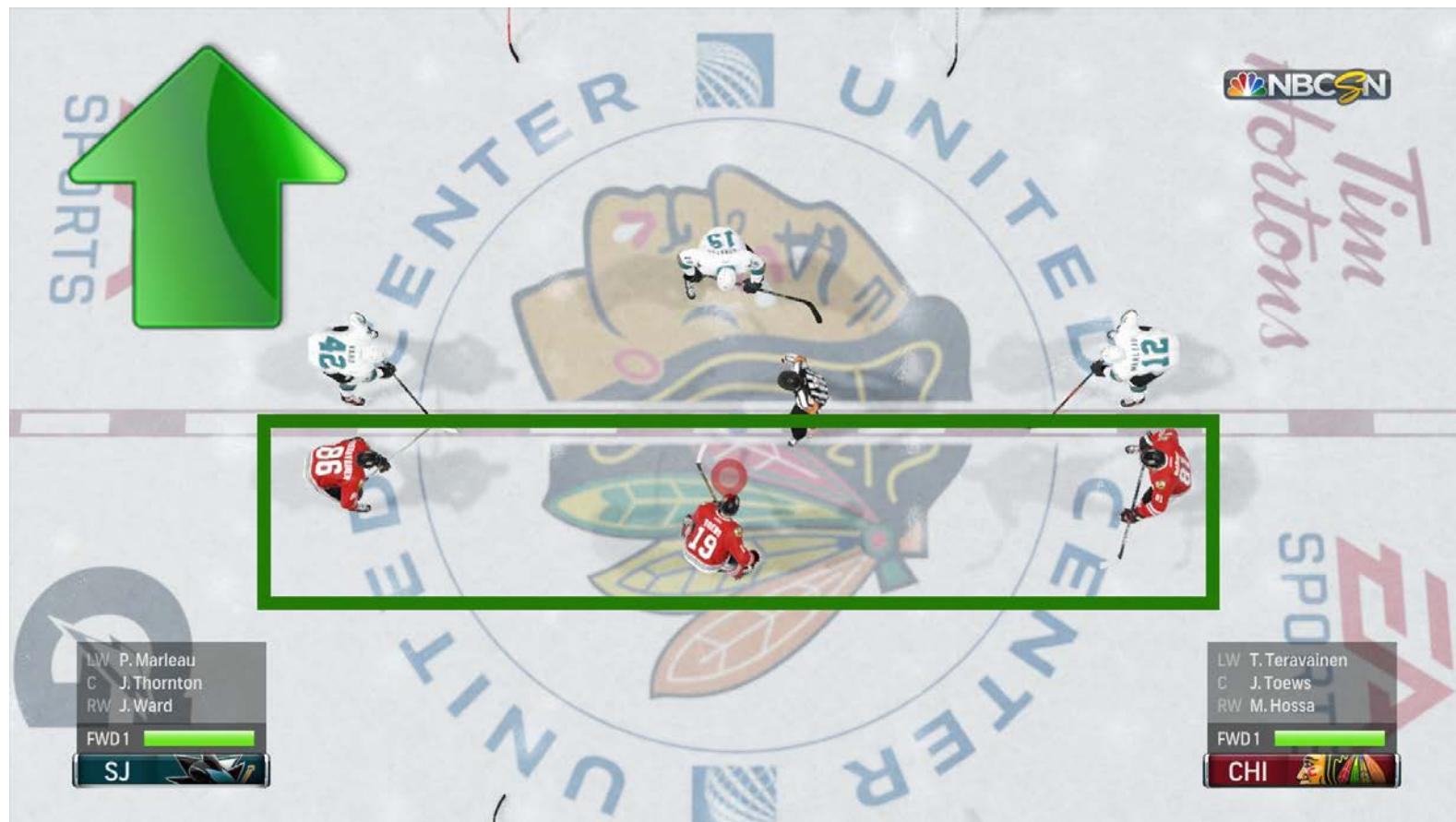


Structural Correlations : Teams



Structural Correlations : Lines

- Goal = 3 pt, assist = 2 pt



Structural Correlations : Lines = Stacking

- At least 1 complete line (3 players per line)
- At least 2 partial lines (at least 2 players per line)

1 complete line constraint

$$3v_i \leq \sum_{p \in L_i} x_{pl}, \quad \forall i \in \{1, \dots, N_L\}$$

$$\sum_{i=1}^{N_L} v_i \geq 1$$

$$v_i \in \{0, 1\}, \quad \forall i \in \{1, \dots, N_L\}.$$

2 partial lines constraint

$$2w_i \leq \sum_{p \in L_i} x_{pl}, \quad \forall i \in \{1, \dots, N_L\}$$

$$\sum_{i=1}^{N_L} w_i \geq 2$$

$$w_i \in \{0, 1\}, \quad \forall i \in \{1, \dots, N_L\}.$$

Structural Correlations : Goalie Against Opposing Players



Structural Correlations : Goalie Against Opposing Players

- No skater against goalie

No skater against goalie constraint

$$6x_{pl} + \sum_{q \in Opponents_p} x_{ql} \leq 6, \quad \forall p \in G$$

Good, but not great chance

Feasible

Line

Team

Line

Goalie

Not
Against



Play many diverse Lineups

- Make sure lineup l has no more than γ players in common with lineups 1 to $l-1$

Diversity constraint

$$\sum_{p=1}^N x_{pk}^* x_{pl} \leq \gamma, k = 1, \dots, l-1$$

Were we able to do it?

NHL \$2K Sniper [\$2,000 Guaranteed]			
STANDINGS	ENTRIES	DETAILS	GAMES
1st	zlisto \$150.00	54.50	PMR 0
3rd	zlisto \$90.00	51.50	PMR 0
9th	zlisto \$30.00	49.50	PMR 0
23rd	zlisto \$18.75	46.00	PMR 0
28th	zlisto \$15.00	45.50	PMR 0
28th	zlisto \$15.00	45.50	PMR 0

NHL \$40K Sniper [\$40,000 Guaranteed]			
STANDINGS	ENTRIES	DETAILS	GAMES
2nd	zlisto \$2,000.00	61.30	PMR 0
21st	zlisto \$50.00	57.30	PMR 0
21st	zlisto \$50.00	57.30	PMR 0
40th	zlisto \$40.00	56.10	PMR 0
42nd	zlisto \$40.00	55.70	PMR 0
81st	zlisto \$10.00	54.10	PMR 0
15th	zlisto \$0.00	50.10	PMR 0

November 15, 2015

November 16, 2015

November 17, 2015

November 23, 2015

NHL \$80K Tuesday Special [\$80,000 Guaranteed]			
STANDINGS	ENTRIES	DETAILS	GAMES
3rd	zlisto \$3,000.00	54.60	PMR 0
6th	zlisto \$1,000.00	52.80	PMR 0
7th	zlisto \$800.00	52.30	PMR 0
10th	zlisto \$600.00	50.60	PMR 0
11th	zlisto \$500.00	50.30	PMR 0
15th	zlisto \$0.00	50.10	PMR 0

NHL \$45K Sniper [\$45,000 Guaranteed]			
STANDINGS	ENTRIES	DETAILS	GAMES
1st	zlisto \$3,000.00	52.60	PMR 0
8th	zlisto \$275.00	49.60	PMR 0
57th	zlisto \$50.00	45.60	PMR 0
57th	zlisto \$50.00	45.60	PMR 0
83rd	zlisto \$40.00	44.60	PMR 0
83rd	zlisto \$0.00	44.60	PMR 0

200 lineups

Policy Change



200 lineups -> 100 lineups

Were we able to continue it?

1st	zlisto	62.50	PMR 0
6th	zlisto	58.80	PMR 0
8th	zlisto	57.40	PMR 0
13th	zlisto	55.80	PMR 0
16th	zlisto	55.30	PMR 0
20th		54.00	



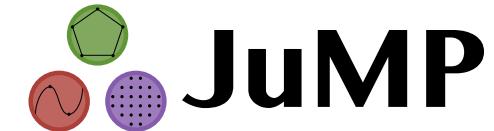
> \$15K

December 12, 2015

100 lineups



How can you do it?



Download Code from Github:

<https://github.com/dscotthunter/Fantasy-Hockey-IP-Code>

<http://arxiv.org/pdf/1604.01455v1.pdf>

Performance Time < 30 Minutes

