

ORF 363/COS 323

Computing and Optimization in the Physical and Social Sciences

Amir Ali Ahmadi

Princeton, ORFE

Lecture 1

What is optimization?

- Roughly, can think of optimization as the science of making the most out of every situation.
- You've probably all done it many times this week:

▪ What courses to take?

- To maximize learning.
- To maximize GPA (?!)
- Courses can't conflict.
- Not before 10AM.
- Professor rating > 4.5.

▪ What furniture to buy?

- To minimize cost.
- To maximize comfort.
- Must fit in your room.
- Must have 3 drawers.
- Not too heavy.

▪ Where to get dinner?

- To minimize cost.
- Less than .5 miles from dormitory.
- Must have ice cream for dessert.
- Sanitation grade > 7.

▪ Common theme:

- You make decisions and choose one of many alternatives.
- You hope to maximize or minimize something (you have an objective).
- You cannot make arbitrary decisions. Life puts constraints on you.

How is this class different from your every-day optimization?

- We'll be learning techniques for dealing with problems that have
 - Thousands (if not millions) of variables
 - Thousands (if not millions) of constraints
- These problems appear every day in the industry, in science, in engineering
- Hopeless to make decisions in your head and with rules of thumb
- Need mathematical techniques that translate into algorithms
 - Algorithms then get implemented on a computer to solve your optimization problem
- We typically model a physical or social scenario with a precise mathematical description
- In this mathematical model, we care about actually finding *the best solution*
- Whenever we can't find the best solution, we would like to know how far off our proposed solution is

Examples of optimization problems

In finance

■ In what proportions to invest in 500 stocks?

- To maximize return.
- To minimize risk.
- No more than 1/5 of your money in any one stock.
- Transactions costs < \$70.
- Return rate > 2%.

In control engineering

■ How to drive an autonomous vehicle from A to B?

- To minimize fuel consumption.
- To minimize travel time.
- Distance to closest obstacle > 2 meters.
- Speed < 40 miles/hr.
- Path needs to be smooth (no sudden changes in direction).

In machine learning

■ How to assign likelihoods to emails being spam?

- To minimize probability of a false positive.
- To penalize overfitting on training set.
- Probability of false negative < .15.
- Misclassification error on training set < 5%.

Computing and Optimization

- This class will give you a broad introduction to
“optimization from a computational viewpoint.”
- Optimization and computing are very close areas of applied mathematics:
 - For a host of major problems in computer science, the best algorithms currently come from the theory of optimization.
 - Conversely, foundational work by computer scientists has led to a shift of focus in optimization theory from “mathematical analysis” to “computational mathematics.”
- Several basic topics in scientific computing (that we’ll cover in this course) are either special cases or fundamental ingredients of more elaborate optimization algorithms:
 - Least squares, root finding, solving linear systems, solving linear inequalities, approximation and fitting, matrix factorizations, conjugate gradients,...



Agenda for today

- Meet your teaching staff
- Get your hands dirty with algorithms
 - Game 1
 - Game 2
- Modelling with a mathematical program
 - Fermat's last theorem!
- Course logistics and expectations

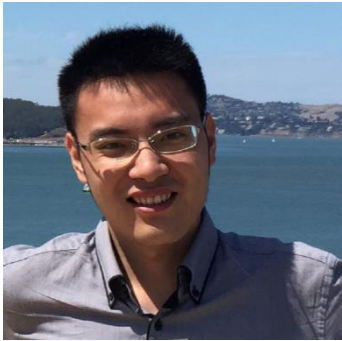
Meet your teaching staff



- **Amir Ali Ahmadi** (Amir Ali, or Amirali, is my first name)
- I am an Assistant Professor at ORFE (since Fall 2014).
- I come here from MIT, EECS, after a fellowship at IBM Research.
- Office hours: **Tuesdays, 8-10 PM, Sherrerd 329. (Overflow room → Sherrerd 125)**
(May move to 6-8 PM some Tuesdays. Will announce in advance.)

<http://aaa.princeton.edu/> a_a_a@p...

Meet your teaching staff



▪ **Zongxi Li**

▪ Office hours: **Mon 4-6,**
Sherrerd 005

▪ [zongxil@p...](mailto:zongxil@princeton.edu)



▪ **Kaizheng Wang**

▪ Office hours: **Mon 6-8,**
Sherrerd 005

▪ [kaizheng@p...](mailto:kaizheng@princeton.edu)



▪ **Bachir El Khadir**

▪ Office hours: **Tue 6-8,**
Sherrerd 005

▪ [bkhadir@p...](mailto:bkhadir@princeton.edu)



▪ **Jeffrey Zhang**

▪ Office hours: **Wed 4-6,**
Sherrerd 107

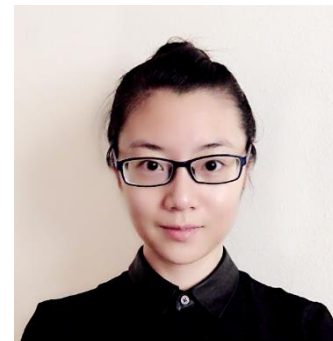
▪ [jeffz@p...](mailto:jeffz@princeton.edu)



▪ **Georgina Hall (honorary TA)**

▪ Office hours: **Wed 6-8,**
Sherrerd 005

▪ [gh4@p...](mailto:gh4@princeton.edu)



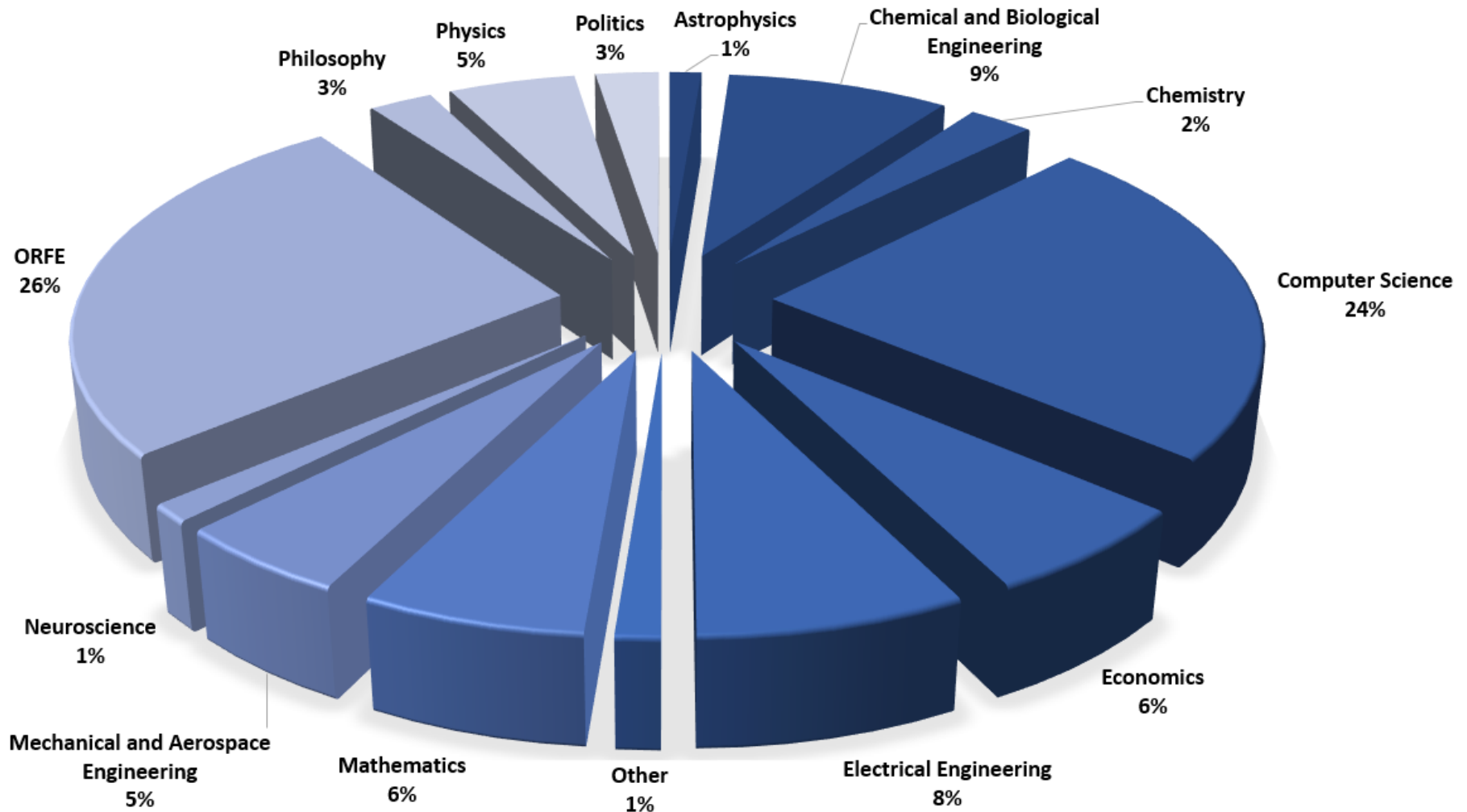
▪ **Jing Ye**

▪ Office hours: **Fri 3-5,**
Sherrerd 107

▪ [jingy@p...](mailto:jingy@princeton.edu)

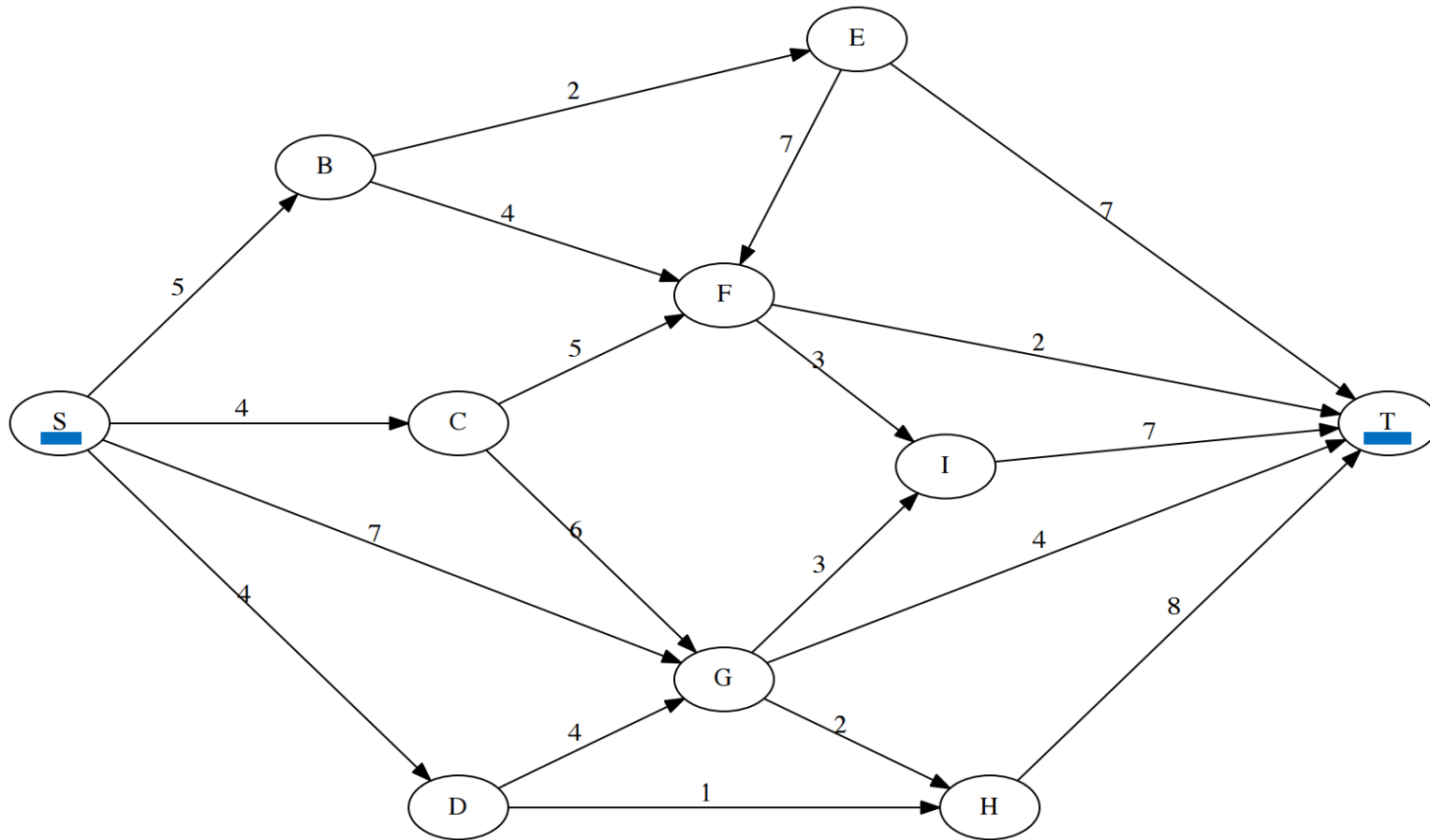
Meet your classmates

ORF 363/COS 323, Fall 2016 (80 registered students)



Let's get to the games!

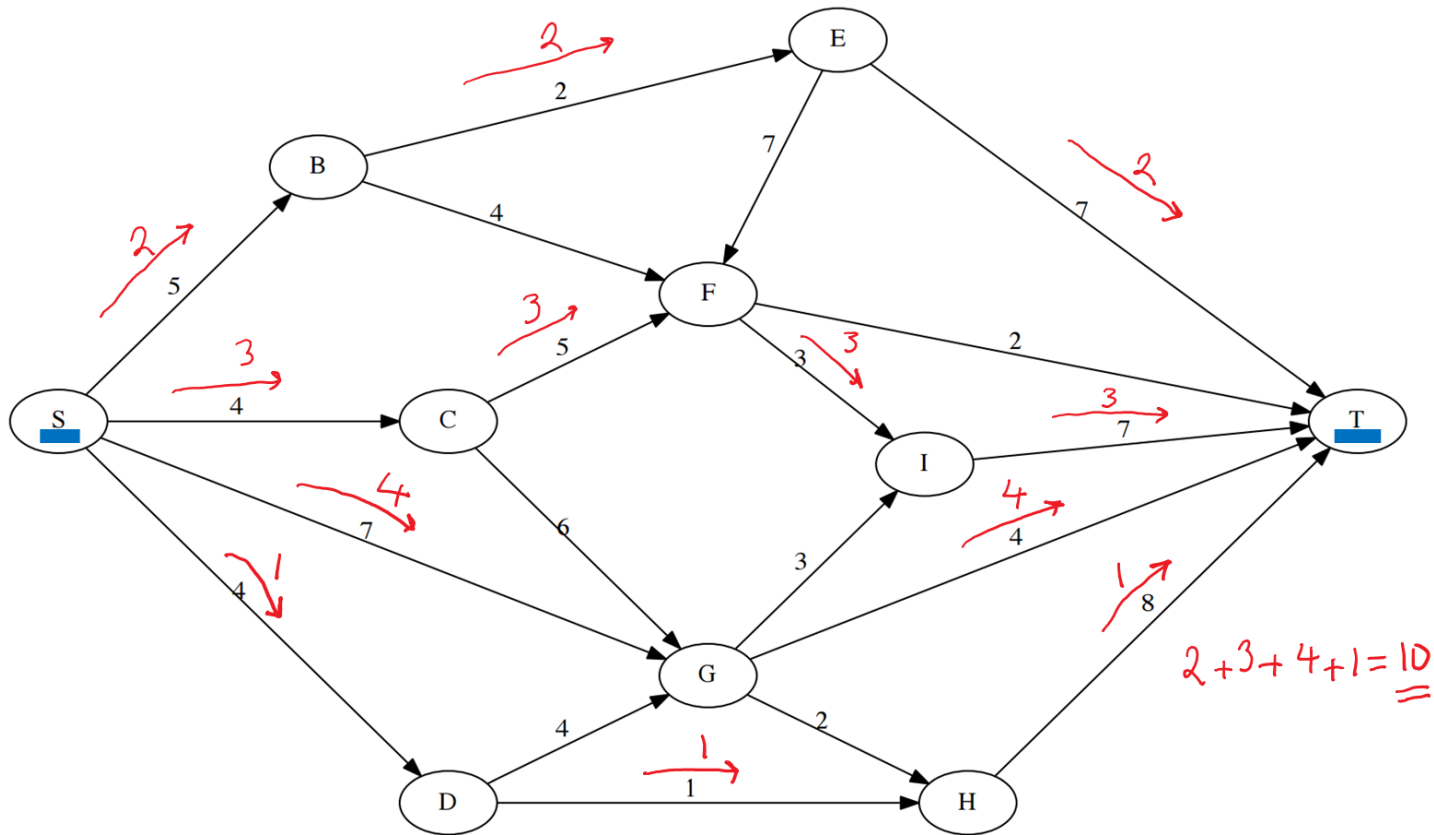
Let's ship some oil together!



■ Rules of the game:

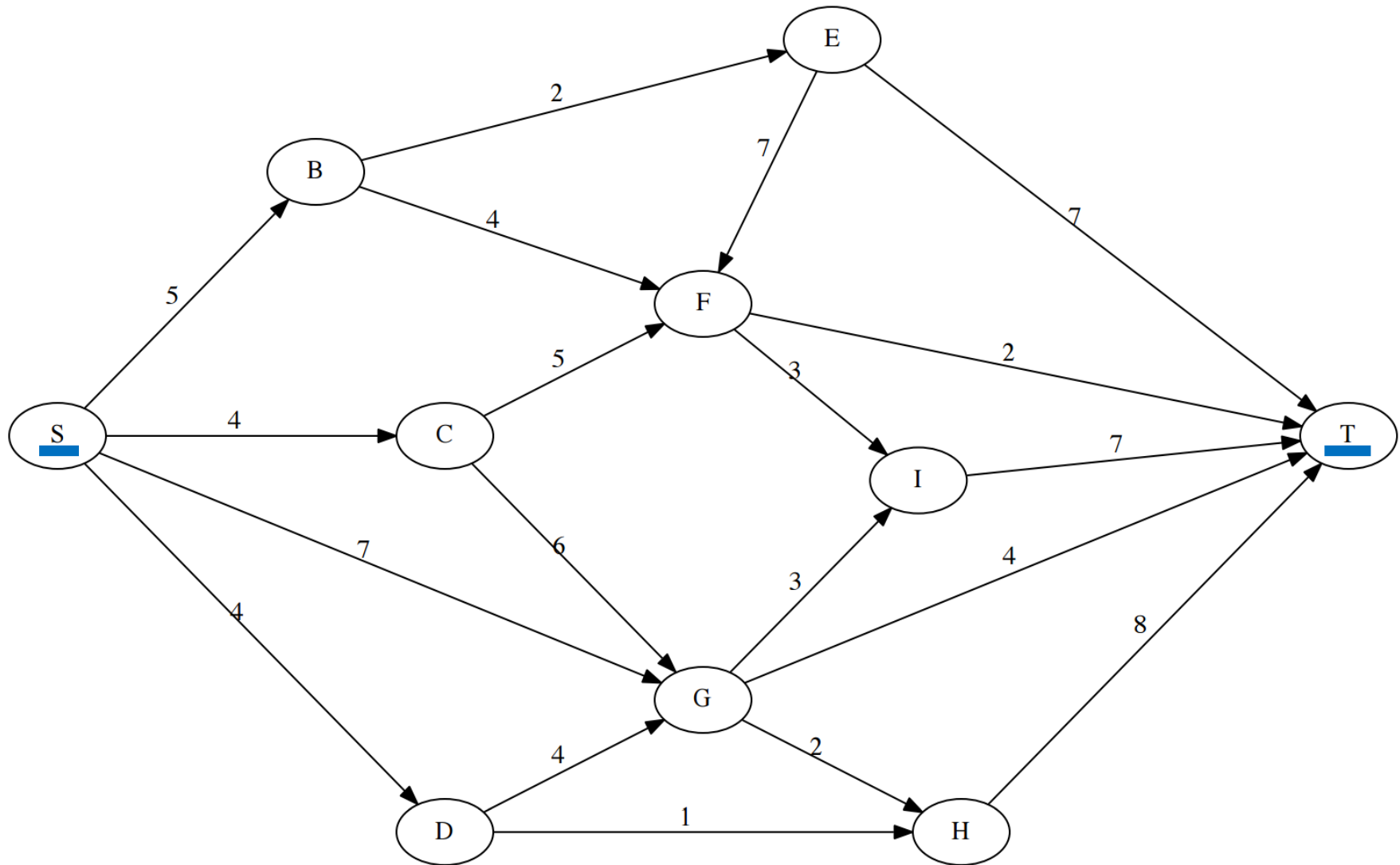
- Cannot exceed capacity on the edges.
- For each node, except for S and T, flow in = flow out (i.e., no storage).
- **Goal:** ship as much oil as you can from S to T.

- Let me start things off for you. Here is a flow with value 10:

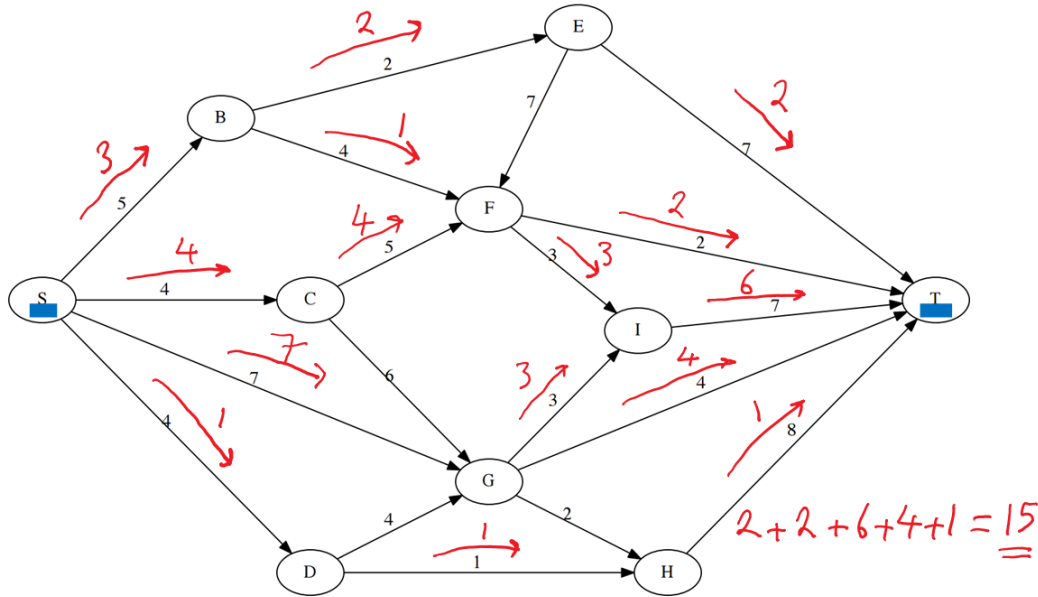


- Can you do better? How much better?
- You all get a copy of this graph on the handout.

You tell me, I draw...

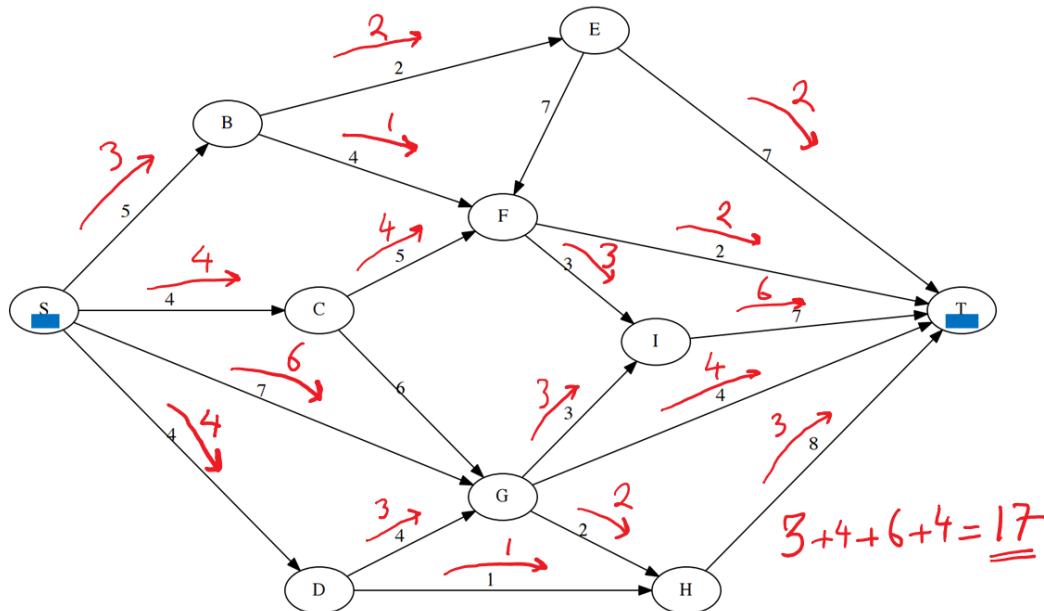


A couple of good attempts



■ Flow of value **15**

■ Can you do better?



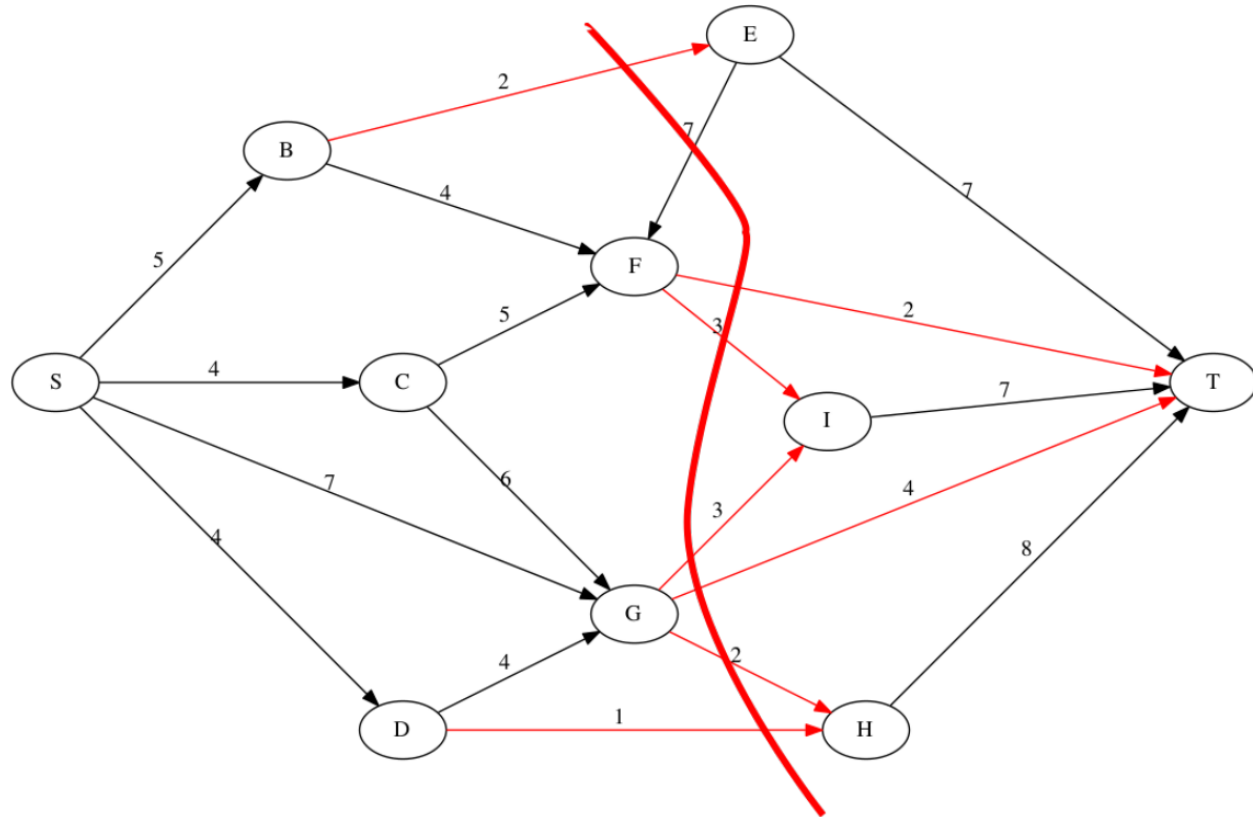
■ Flow of value **17**

■ Can you do better?

■ How can you prove that it's impossible to do better?

17 is the best possible!

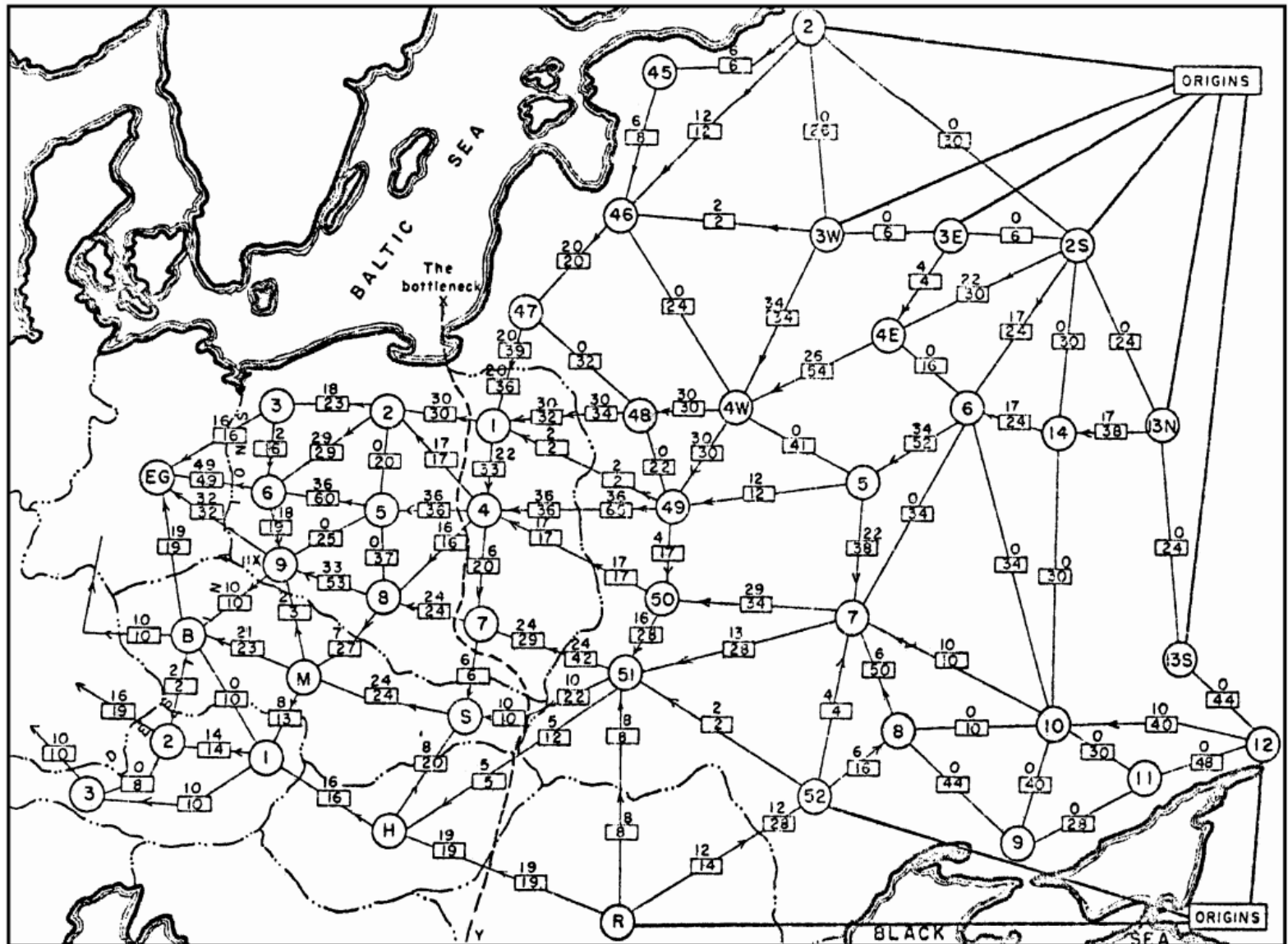
- Proof by magic:



- The rabbit is the red “cut”!
- Any flow from S to T must cross the red curve.
- So it can have value at most 17.

- And here is the magic: such a proof is *always* possible! 15

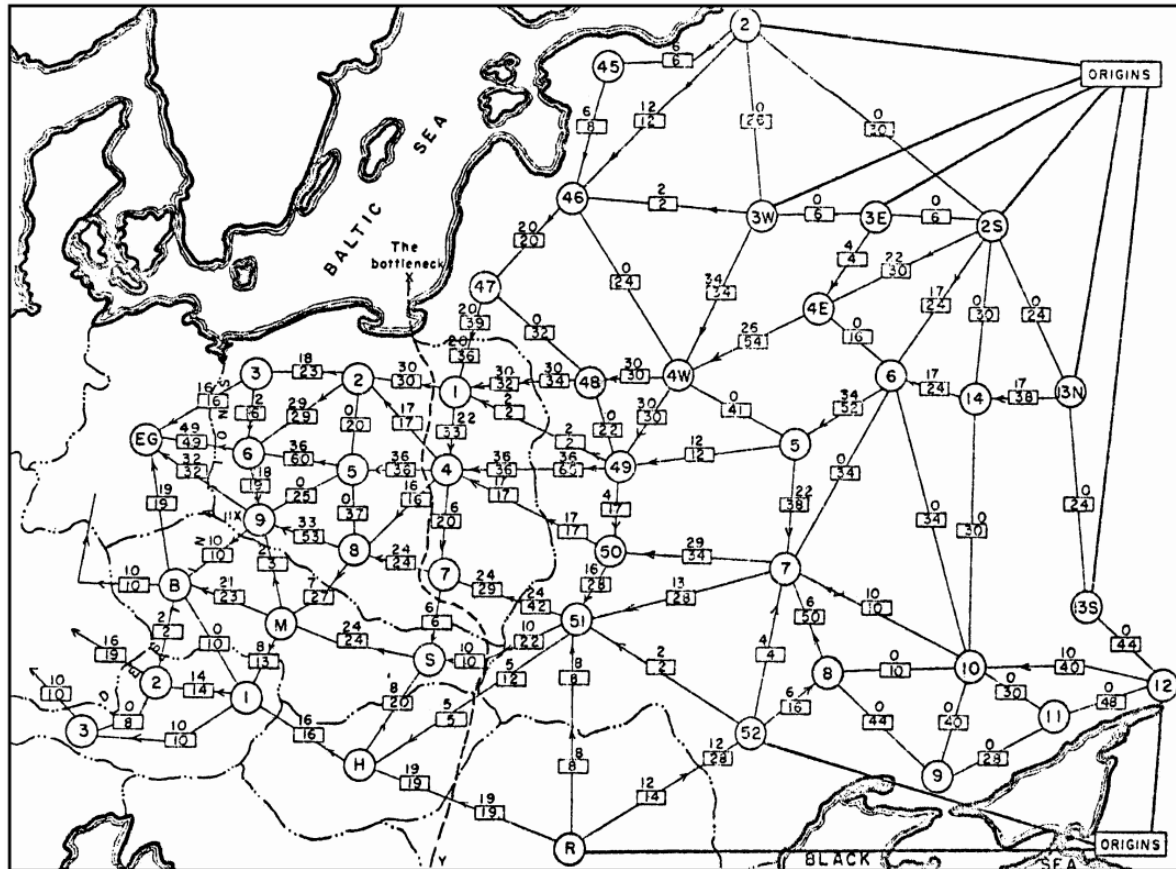
Let's try a more realistic graph



■ How long do you think an optimization solver would take (on my laptop) to find the best solution here?

■ How many lines of code do you think you have to write for it?

■ How would someone who hasn't seen optimization approach this?



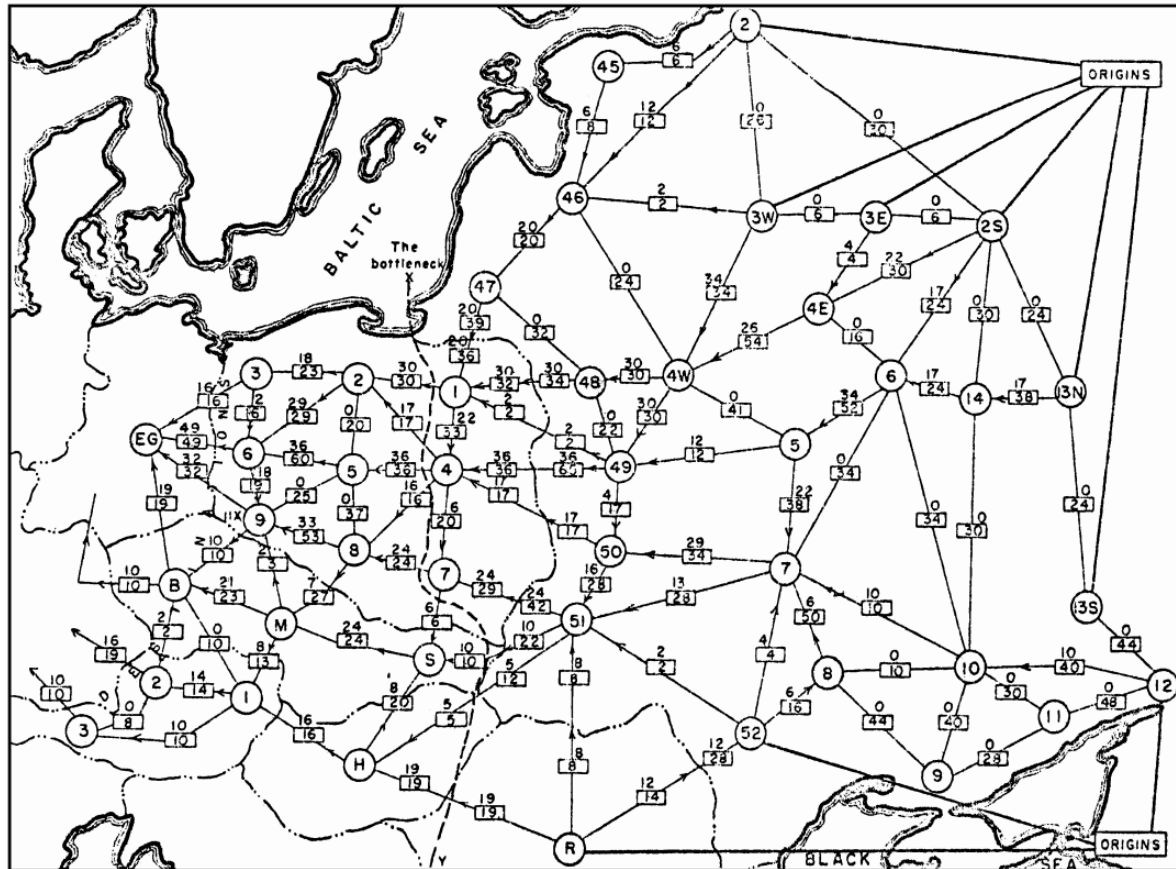
■ Trial and error?

■ Push a little flow here, a little there...

■ Do you think they are likely to find the best solution?

A bit of history behind this map

- From a secret report by Harris and Ross (1955) written for the Air Force.
- Railway network of the Western Soviet Union going to Eastern Europe.
- Declassified in 1999.
- Look at the min-cut on the map (called the “bottleneck”)!
- There are 44 vertices, 105 edges, and the max flow is 163K.



- Harris and Ross gave a heuristic which happened to solve the problem optimally in this case.
- Later that year (1955), the famous Ford-Fulkerson algorithm came out of the RAND corporation. The algorithm always finds the best solution (for rational edge costs).

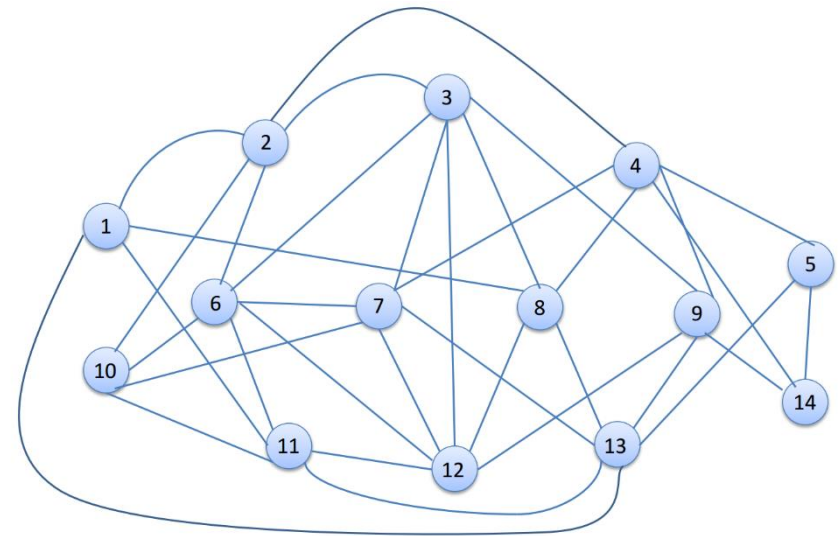
Let's look at a second problem

...and tell me which one you
thought was easier

Two finals in one day? No thanks.

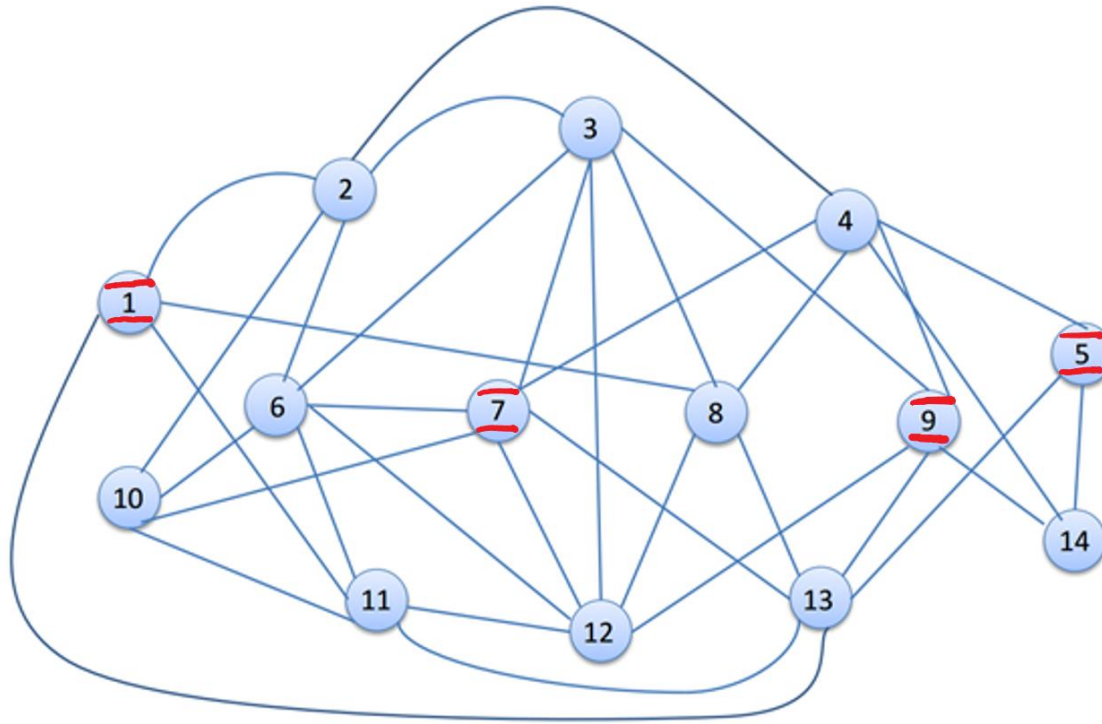
- The department chair at ORFE would like to schedule the final exams for 14 undergraduate courses offered this semester.
- He wants to have as many exams as possible on the same day, so everyone gets done quickly and goes on vacation.
- There is just one constraint:
No student should have >1 exam on the same day.

- The nodes of this graph are the 14 courses.
- There is an edge between two nodes if and only if there is at least one student who is taking both courses.
- If we want to schedule as many exams as possible on the same day, what are we looking for in this graph?



- The largest collection of nodes such that no two nodes share an edge.

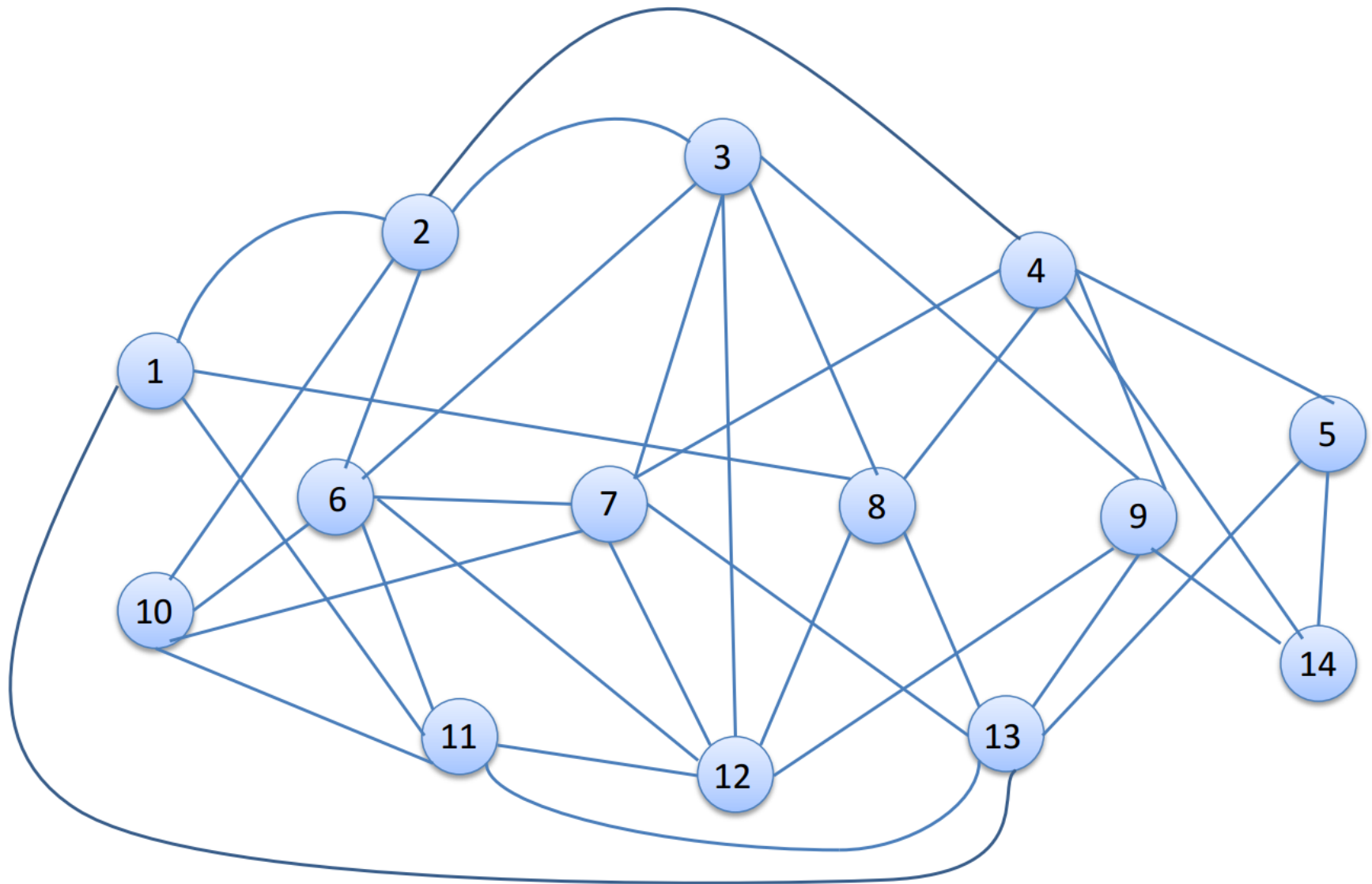
- Let me start things off for you. Here is 4 concurrent final exams:



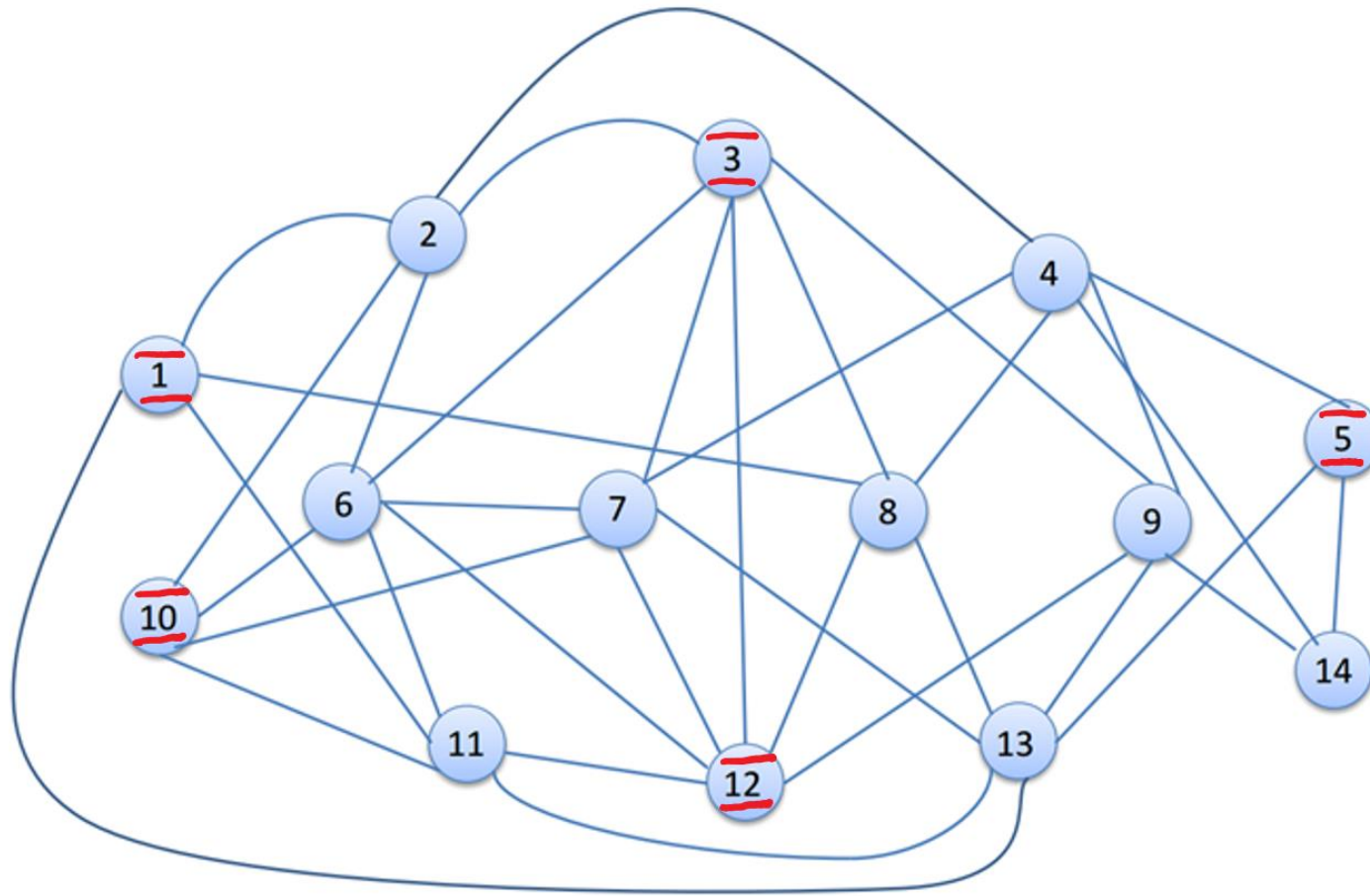
- Can you do better?
- How much better?
- You all get a copy of this graph on the handout.

You have 7 minutes!

You tell me, I draw...



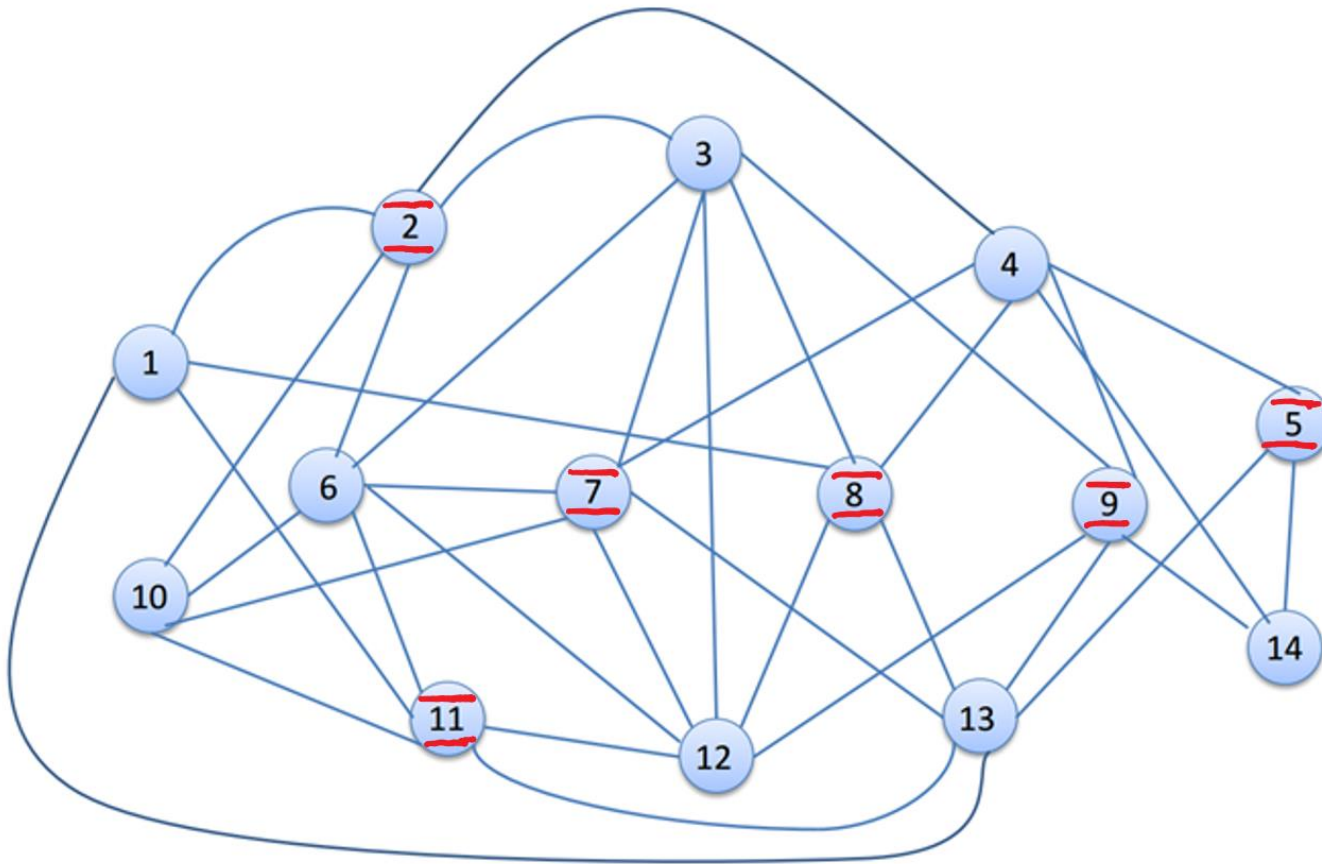
A couple of good attempts



5 exams

■ Can you do better?

A couple of good attempts



6 exams

- Tired of trying?
- Is this the best possible?

6 is the best possible!

- Proof by magic?



- Unfortunately not ☹

- No magician in the world has pulled out such a rabbit to this day! (By this we mean a rabbit that would work on *all* graphs.)

- Of course there is always a proof:

- Try all possible subsets of 7 nodes.

- There are 3432 of them.

- Observe that none of them work.

- But this is no magic. It impresses nobody. We want a “short” proof. (We will formalize what this means.) Like the one in our max-flow example.

- Let’s appreciate this further...

Let's try another graph

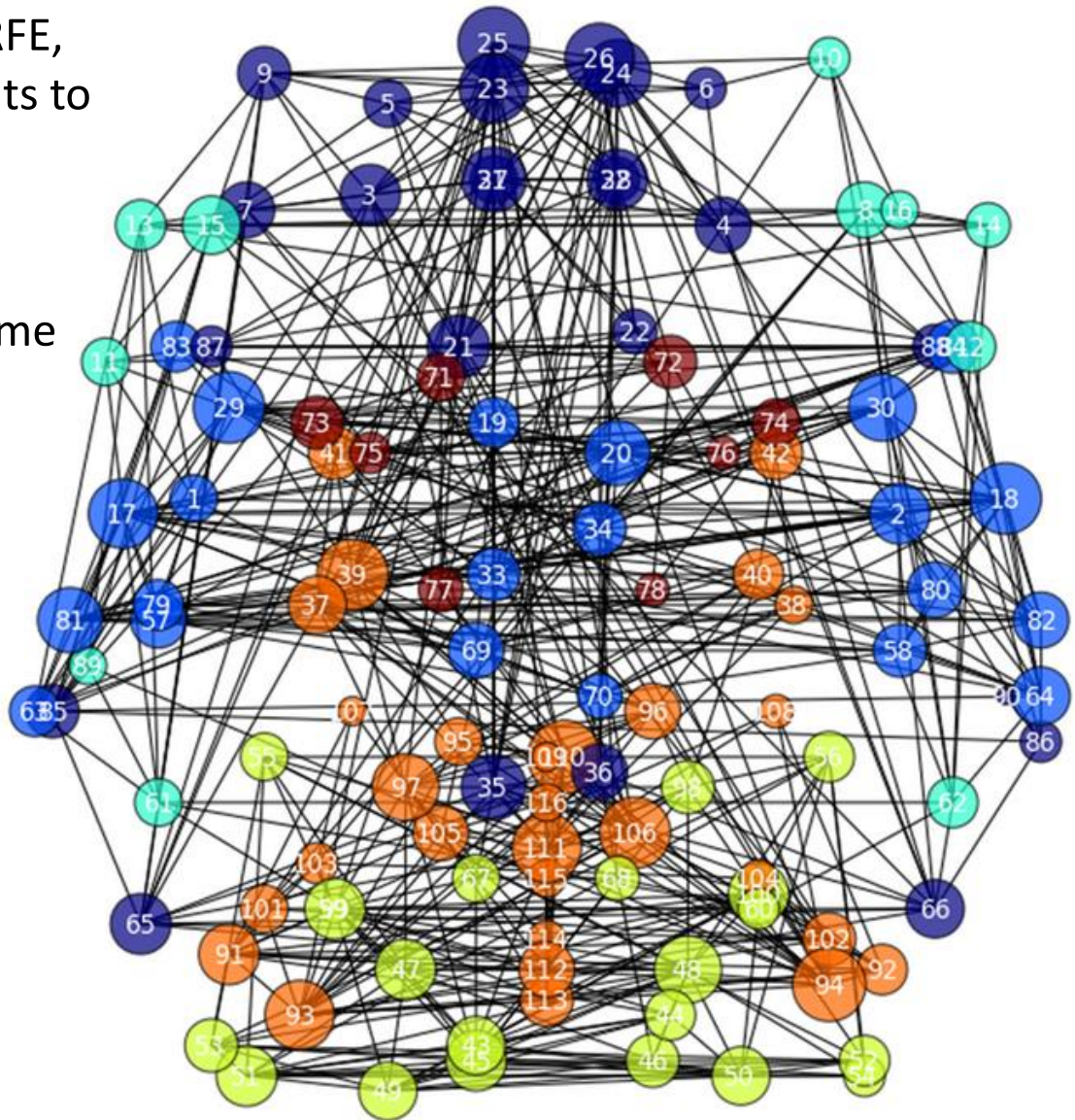
- Encouraged by the success of ORFE, now the Dean of Engineering wants to do the same for 115 SEAS courses.

- How many final exams on the same day are possible? Can you do 17?

 - You have 7 minutes! ;)

- Want to try out all possibilities for 17 exams?

- There are over
8000000000000000000000000
of them!



But there is some good news

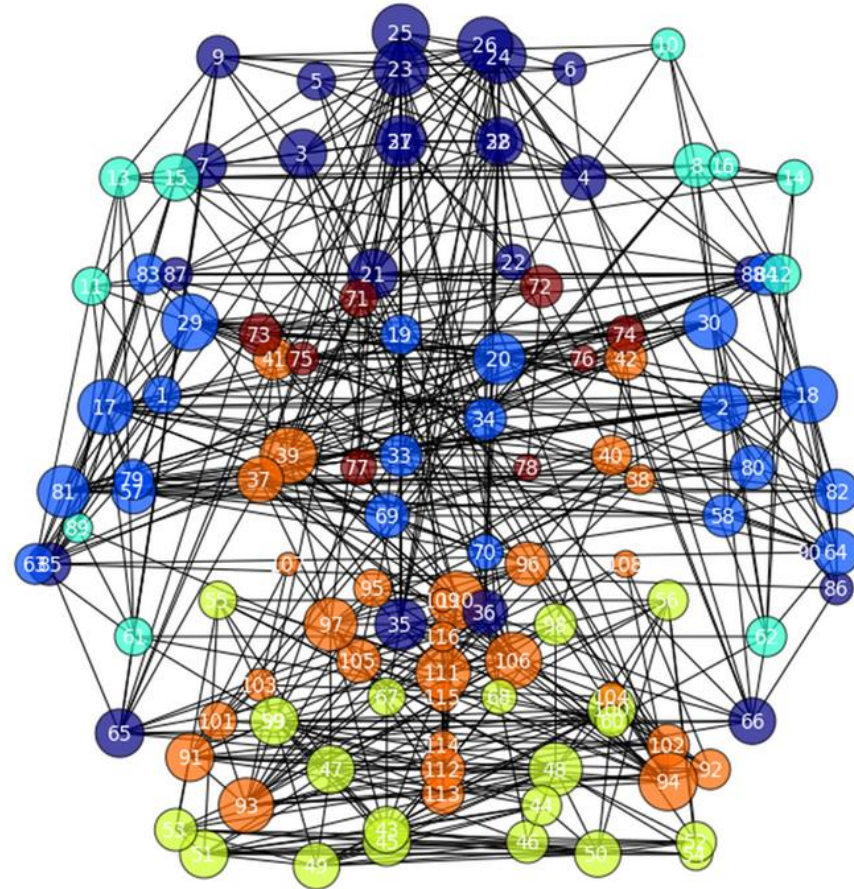
- Even though finding the best solution always may be too much to hope for, techniques from optimization (and in particular from the area of *convex optimization*) often allow us to find high-quality solutions with performance guarantees.

- For example, an optimization algorithm may quickly find 16 concurrent exams for you.

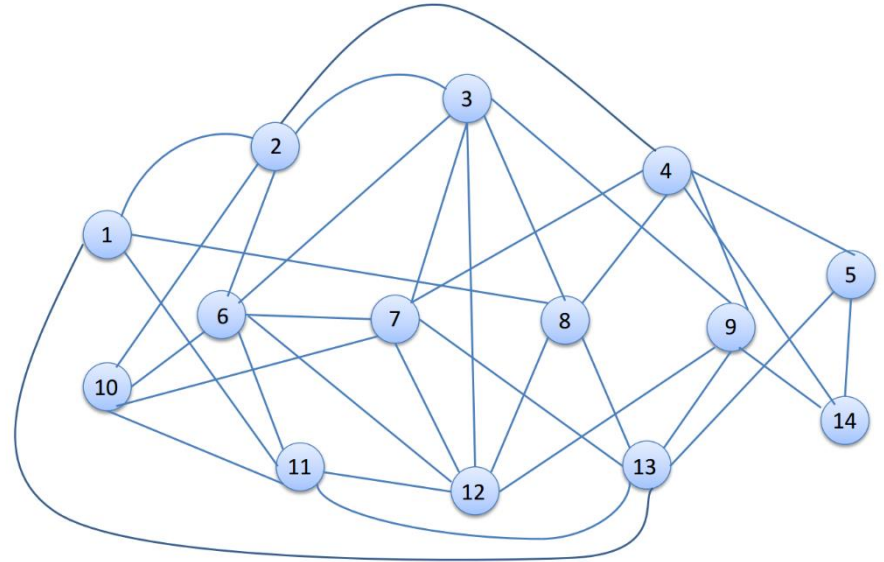
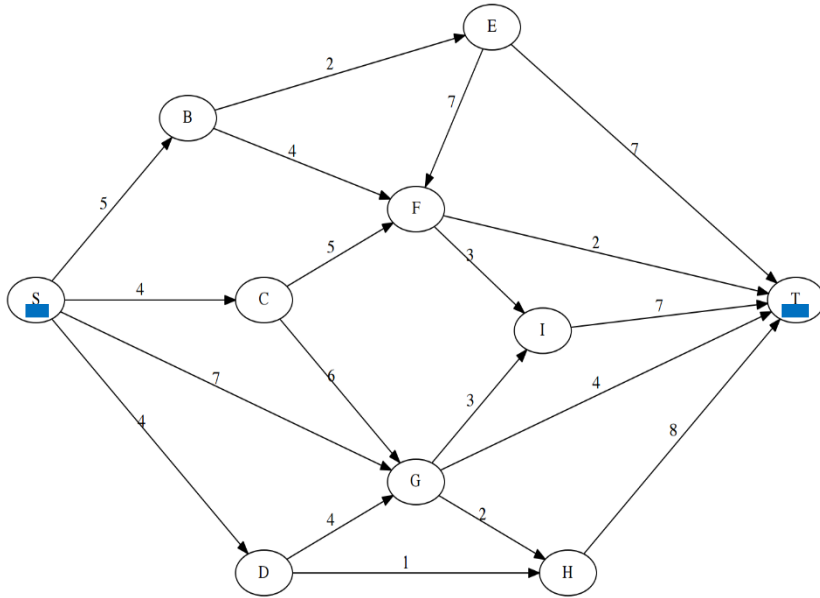
- You really want to know if 17 is impossible. Instead, another optimization algorithm (or sometimes the same one) tells you that 19 is impossible.

- This is very useful information! You know you got 16, and no one can do better than 19.

- We will see a lot of convex optimization in this class!



Which of the two problems was harder for you?

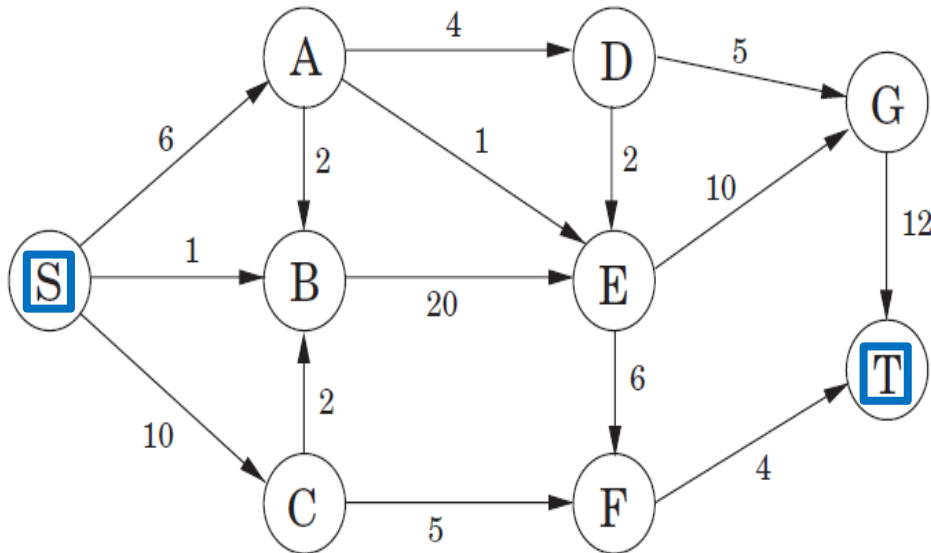


- Not always obvious. A lot of research in optimization and computer science goes into distinguishing the “tractable” problems from the “intractable” ones.
- The two brain teasers actually just gave you a taste of the **P vs. NP** problem. (If you have not heard about this, that’s OK. You will soon.)
- The first problem we can solve efficiently (in “polynomial time”).
- The second problem: no one knows. If you do, you literally get \$1M!
 - More importantly, your algorithm immediately translates to an efficient algorithm for thousands of other problems no one knows how to solve.



Modelling problems as a mathematical program

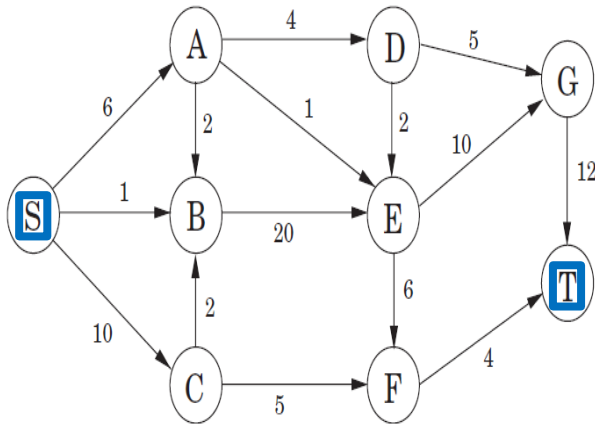
Let's revisit our first game



- What were your decision variables?
- What were your constraints?
- What was your objective function?

▪Rules of the game:

- Cannot exceed capacity on the edges.
- For each node, except for S and T, flow in = flow out (i.e., no storage).
- Goal:** ship as much oil as you can from S to T.



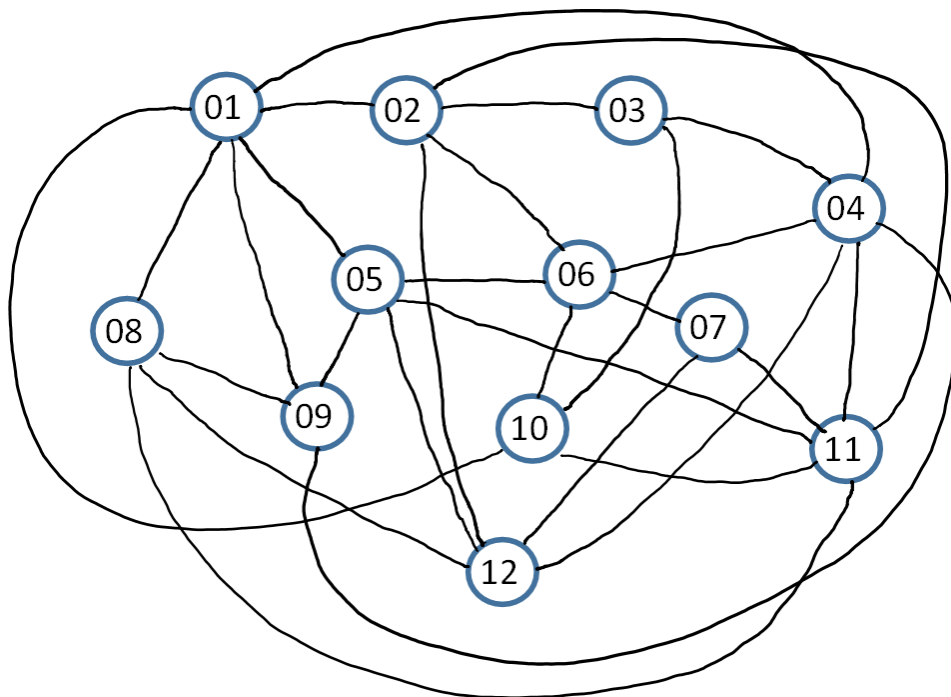
$x_{SA}, x_{AD}, x_{BE}, \dots, x_{GT}$ ← Decision variables

max. $x_{SA} + x_{SB} + x_{SC}$ ← Objective function

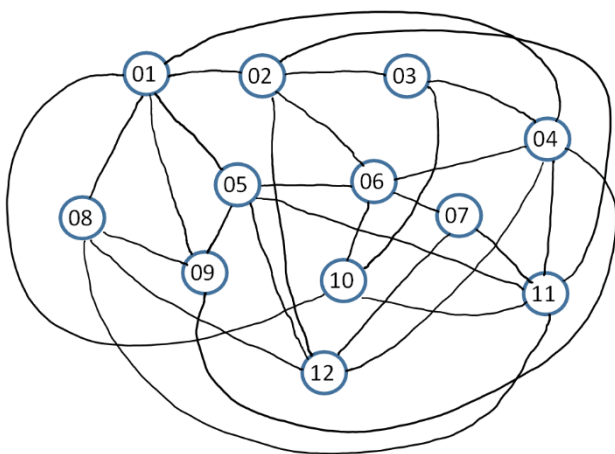
s.t.

- $x_{SA}, x_{AD}, x_{BE}, \dots, x_{GT} \geq 0$
 - $x_{SA} \leq 6, x_{AB} \leq 2, x_{EG} \leq 10, \dots, x_{GT} \leq 12$
 - $\begin{cases} x_{SA} = x_{AD} + x_{AB} + x_{AE} \\ x_{SC} = x_{CB} + x_{CF} \\ \vdots \\ x_{CF} + x_{EF} = x_{FT} \end{cases}$
- ← Constraints

Let's revisit our second game



- What were your decision variables?
- What were your constraints?
- What was your objective function?



$x_1, x_2, \dots, x_{12}$ ← Decision variables

max. $x_1 + x_2 + \dots + x_{12}$ ← Objective function

s.t.

o $x_i (1 - x_i) = 0, i = 1, \dots, 12$

o $\left[\begin{array}{l} x_1 + x_2 \leq 1 \\ x_1 + x_8 \leq 1 \\ x_4 + x_6 \leq 1 \\ \vdots \\ x_{12} + x_8 \leq 1 \end{array} \right. \quad \text{(one per edge)}$

← Constraints

Why one hard and one easy? How can you tell?

$$x_{SA}, x_{AD}, x_{BE}, \dots, x_{GT}$$

$$\max. \quad x_{SA} + x_{SB} + x_{SC}$$

s.t.

$$\circ \quad x_{SA}, x_{AD}, x_{BE}, \dots, x_{GT} \geq 0$$

$$\circ \quad x_{SA} \leq 6, x_{AB} \leq 2, x_{EG} \leq 10, \dots, x_{GT} \leq 12$$

$$\circ \quad \begin{cases} x_{SA} = x_{AD} + x_{AB} + x_{AE} \\ x_{SC} = x_{CB} + x_{CF} \\ \vdots \\ x_{CF} + x_{EF} = x_{FT} \end{cases}$$

$$x_1, x_2, \dots, x_{12}$$

$$\max. \quad x_1 + x_2 + \dots + x_{12}$$

s.t.

$$\circ \quad x_i (1 - x_i) = 0, \quad i = 1, \dots, 12$$

$$\circ \quad \begin{cases} x_1 + x_2 \leq 1 \\ x_1 + x_8 \leq 1 \\ x_4 + x_6 \leq 1 \\ \vdots \\ x_{12} + x_8 \leq 1 \end{cases} \quad (\text{one per edge})$$

■ **Caution:** just because we can write something as a mathematical program, it doesn't mean we can solve it.

Fermat's Last Theorem

- Can you give me three positive integers x, y, z such that

$$x^2 + y^2 = z^2?$$

- Sure:

(3, 4, 5)	(5, 12, 13)	(8, 15, 17)	(7, 24, 25)
(20, 21, 29)	(12, 35, 37)	(9, 40, 41)	(28, 45, 53)

And there are infinitely many more...

- How about $x^3 + y^3 = z^3?$

- How about $x^4 + y^4 = z^4?$

- How about $x^5 + y^5 = z^5?$

Fermat's Last Theorem

- Fermat's conjecture (1637):

For $n \geq 3$, the equation $x^n + y^n = z^n$ has no solution over positive integers.

- Proved in 1994 (357 years later!) by Andrew Wiles.

(Was on the faculty in our math department until a few years ago.)



Arithmeticon Liber II.

61

interuallum numerorum 2. minor autem
1 N. atque ideo maior 1 N. + 2. Oportet
itaque 4 N. + 4. triplos esse ad 2. & ad-
huc superaddere 10. Ter igitur 2. adscit-
is vnitatibus 10. æquatur 4 N. + 4. &
fit 1 N. 3. Erit ergo minor 3. maior 5. &
satisfaciunt quæstioni.

εἰς ἵνους. οἱ ἀσπυ μίλλον ἵσους εἰς ἵνους μ. β. δὲ
 σσι ἀσπυ ἀελυθινὸν δ' ὑπὸ μῆδους δ' ἀπὸ πλάσσειας
 ἡ) μ. β. εἰς ἵνους ὑπὸ πλάσσειας μ. γ. τρις ἀσπυ
 μῆδους β. μ. γ. μ. δ. ἵσους ὅσους εἰς μ. δ.
 δ. εἰς ἵνους οἱ ἀελυθινὸν μ. γ. ἵσους οἱ μὲν ἵσους
 σσι μ. γ. οἱ δὲ μίλλον μ. γ. εἰς μ. γ. πῶς τῶν
 ποσὺν α.

IN QVAESTIONEM VII.

CONDITIONIS appositæ eadem ratio est quæ & appositæ præcedenti quæstioni, nil enim aliud requirit quàm ut quadratus intervalli numerorum sit minor intervallo quadratorum, & Canones iidem hic etiam locum habebunt, ut manifestum est.

QVÆSTIO VIII.

PROPOSITUM quadratum diuidere in duos quadratos. Imperatum fit vt 16. diuidatur in duos quadratos. Ponatur primus 1 Q. Oportet igitur 16 — 1 Q. equalis esse quadrato. Fingo quadratum a numeris quotquot libuerit, cum defectu totum vitatum quod continet latus ipsius 16. effo 2 1 N. — 4. Ipe igitur quadratus erit 4 Q. + 16. — 1 N. hæc æquabuntur vtriusque 16 — 1 Q. Communis adiciatur vtriusque defectus, & a filiisq. auferatur familia, sient 5 Q. æquales 16 N. & filii 1 N. ¶ Erig igitur quadratorum 5 alter vero 5 & vtriusque summa est 7 seu 16. & vtrique quadratus est.

Τὸν διαπορεύσαντα γαμῶσι διελὺν εἰς
 δύο τετραγώνια. ἐπιτετραχὲν δὲ τὸ
 διελὺν εἰς δύο τετραγώνια, καὶ ἐπιδόξω
 σφιστος διαικῶσι μᾶς, δὴτοις ἀεὶ μῆ-
 ρος τὸν λήγει διαικῶσι μᾶς ἵνα τὸ
 τετραγώνιον, πλάτος δὲ τετραγώνιον εἰς δύο
 δὴ πῶς λέγει πῶς τὸν μὲν δὴ τὸν εἰς
 μὲν πλάτος, ἵνα εἰς β' λέγει μὲν δ', αὐτὸ
 ἀεὶ δὲ πλάτος τὸν διαικῶσι μᾶς δ' εἰς
 λέγει εἰς γ'. ταῦτα τὴν μῆσαν εἰς λέγει
 διαικῶσι μᾶς, κοπὴ σφαιρικῶν δὲ λήγει
 τὸν δὴ οὐκ ὀρθὸν ὄμμα. διαικῶσι ἀεὶ εἰς
 αὐτὸν μῆ- γ' ἵνα τὸν ἀεὶ τὸν δ' πῶς
 τὸν. ἵνα εἰς δὴ τὸν εἰκοσάτονον, δὲ δὴ
 εἰκοσάτον πλάτος, εἰ δὲ δύο σφαιρικῶσι πῶς
 εἰκοσάτον πλάτος.

OBSERVATIO DOMINI PETRI DE FERMAT.

Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos
& generaliter nullam in infinitum ultra quadratum potestatem in duos eius-
dem nominis fas est dividere cuius rei demonstrationem mirabilem sane detexi.
Hanc marginis exigentia non caperet.

QVÆSTIO IX.

RURSUS oporteat quadratum 16 dividere in duos quadratos. Ponatur rursus primi lateris : N. alterius verò quocunque numerorum cum defectu tot vnitatur, quot constat latus diuidendi. Esto itaque 2 N. — 4. erunt quadrati, hic quidem 1 Q. ille vero 4 Q. + 16. — 16 N. Cæterum volo vtrumque simul æquari vnitatibus 16. Igitur 5 Q. + 16. — 16 N. æquatur vnitatibus 16. & fit 1 N. $\frac{7}{4}$ erit

ΕΣΤΩ δὴ πάλιν τὸς ἐς παραχρᾶτον δι-
 λείν ἐς δύο τετραγώνους τεταθῆναι καὶ
 ἢ τὸ πρῶτον πλῆθει ἐς ἑξῆς, ἢ τὴν ἰσότη-
 τιν ὅταν διέλθῃται καὶ ὅταν ἐξῇ ἡ τῆς δια-
 μέτρος πλάτος. ἔστω δὴ ἐς δύο λείπει μὲν
 ἰσοῦται· οὐ τετραγώνους ἐς ἑξῆς διαμέτρον
 ἐς δὲ διαμέτρον ἐς μὲν λείπει ἐς τὴν βί-
 βουται τὸς δύο λαμβάνει συντάξις. (ἔστω μὲν
 ἐς διαμέτρον ἀπὸ αὐτῶν ἐς μὲν λείπει ἐς τὴν
 μὲν ἐς καὶ γίνονται οὐ ἀριθμοὺς ἐς πνεύματα.)

H iii

Fermat's Last Theorem

■ Fermat's conjecture (1637):

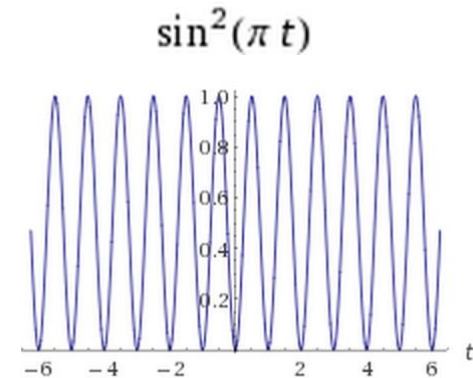
For $n \geq 3$, the equation $x^n + y^n = z^n$ has no solution over positive integers.

■ Consider the following optimization problem (mathematical program):

$$\min_{x, y, z, n} (x^n + y^n - z^n)^2$$

$$\text{s.t. } x \geq 1, y \geq 1, z \geq 1, n \geq 3,$$

$$\sin^2 \pi n + \sin^2 \pi x + \sin^2 \pi y + \sin^2 \pi z = 0.$$



■ Innocent-looking optimization problem: 4 variables, 5 constraints.

■ If you could show the optimal value is non-zero, you would prove Fermat's conjecture!

Course objectives

- The skills I hope you acquire:

- Ability to view your own field through the lens of optimization and computation
 - To help you, we'll draw applications from operations research, statistics, finance, machine learning, engineering, ...
- Learn about several topics in scientific computing
- More mathematical maturity and ability for rigorous reasoning
 - There will be some proofs in lecture. Easier ones on homework.
- Enhance your coding abilities (nothing too fancy, simple MATLAB)
 - There will be a MATLAB component on every homework and on the take-home final.
- Ability to recognize hard and easy optimization problems
- Ability to use optimization software
 - Understand the algorithms behind the software for some easier subclass of problems.

Things you need to download

- Right away:

MATLAB

<http://www.princeton.edu/software/licenses/software/matlab/>

- In the next week or two (will appear on HW#2 or #3):

CVX

<http://cvxr.com/cvx/>