Welcome to 4820/5820

course goals: techniques for designing algorithms

- greedy
- dynamic programming

prelim I

- etc

some problems hard

- NP-completeness
- computability

use of hardness crypto

alg for hard problems

4820/5820 Logistics

Prerequisites 2110 or 2112 & 3110 or A- in other two
2800
data structures, coding in Java or Python
proofs & probability

Section plans
mandatory, practice problems & quiz (on previous hw)

Homework schedule Friday → Friday

Collaboration great, but write solution on your own

More 4820/5820 Logistics

Poll everywhere

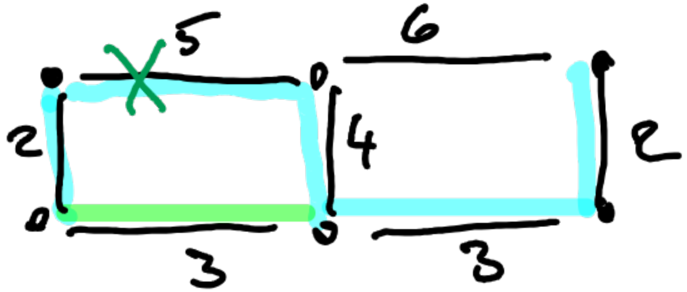
Exams: prelim I Thursday, Feb ~~26~~ $\boxed{12}$
prelim II, Tuesday March 24
final tbd

Office hours starting Friday

All info at our Web page: https://www.cs.cornell.edu/courses/cs4820/2026sp/

Topic 1: Greedy algorithms

- Problem today: Connected graph of minimum cost



given graph $G=(V,E)$

edges $c_e > 0$ cost $e \in E$

- Example

total cost: 16

optimal: $2+3+4+3+2=14$

problem: find minimum cost subgraph

Claim: optimal solution is a tree, i.e. no cycle

Minimum cost spanning tree (MST)

- Basic properties : solution is a tree

proof: contradiction

suppose contains a cycle

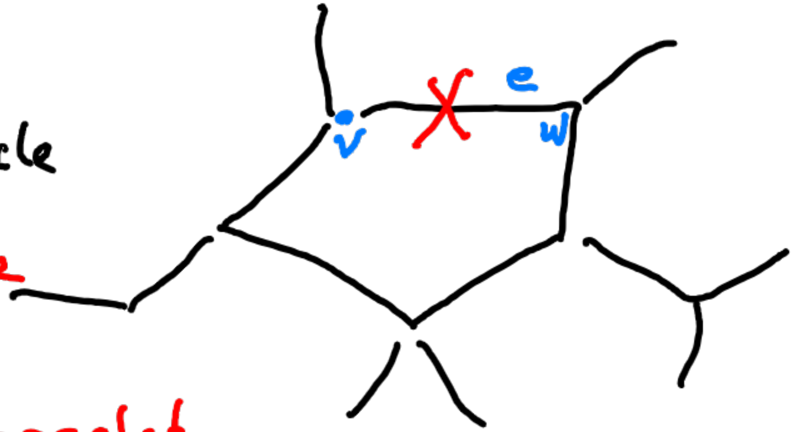
delete any edge from cycle

1. cheaper
2. graph remains connected

as two ends of e connected

by rest of cycle

$e = (v, w)$



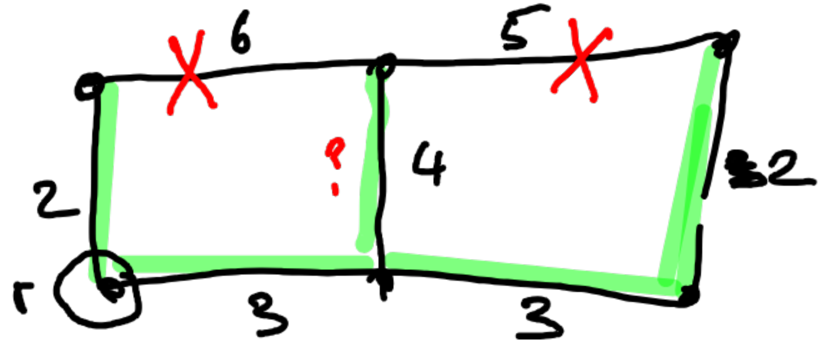
Greedy algorithms for MST

greedy = myopic choices
& no revision

1. cheapest first
   order by increasing cost
   add edges in this order
   unless they form a cycle

prove today

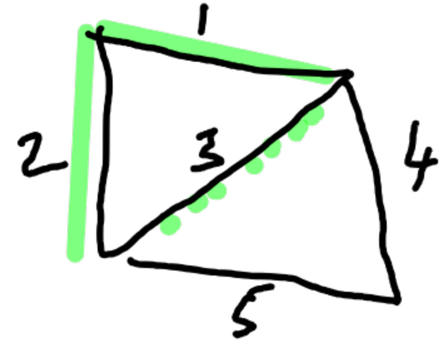
Kruskal



2. select a root
   add edge connects to r
   to new node the cheapest
   way

Prim

3. order edges by decreasing
   cost & remove them unless
   removal disconnects graph



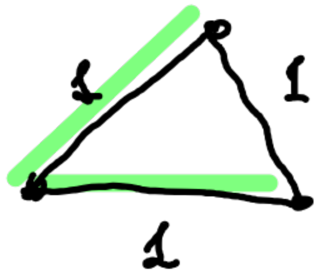
Proving correctness

Kruskal

Proof technique: exchange argument:

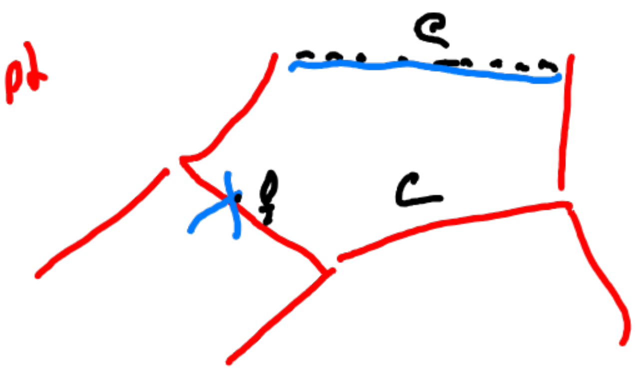
suppose not true: Optimum not same as Kruskal

take optimal solution that agrees with Kruskal with many edges as possible *



Kruskal

Opt



consider $\boxed{\text{first edge } e}$

Kruskal took e is not in Opt

- adding e closes cycle C

Observe: C contains an edge $f \in C$ that was not included before e in Kruskal's solution & e is first time the two solutions differ.

$\Longrightarrow c_f \geq c_e$ *

consider $\boxed{\text{Opt-tree} + e - f}$ (swap e for f

- result: new tree
  - no more expensive * ⇒ another Opt
  - has more shared edges
    contradicting *

Need also running time!