W, 11/20/19

Fall'19 CSCE 629

Analysis of Algorithms

Fang Song
Texas A&M U

Lecture 32

- Linear programming relaxation
- Randomized algorithms

Recall: approximating vertex cover by LP relaxation

(ILP Π) Min $\sum_{i=1}^{n} x_i$
Subject to:

$$x_i + x_j \geq 1, \quad \forall (i,j) \in E$$
$$x_i \in \{0,1\}, \quad \forall i \in V$$

➡

(LP Σ) Min $\sum_{i=1}^{n} x_i$
Subject to:

$$x_i + x_j \geq 1, \quad \forall (i,j) \in E$$
$$0 \leq x_i \leq 1, \quad \forall i \in V$$

?

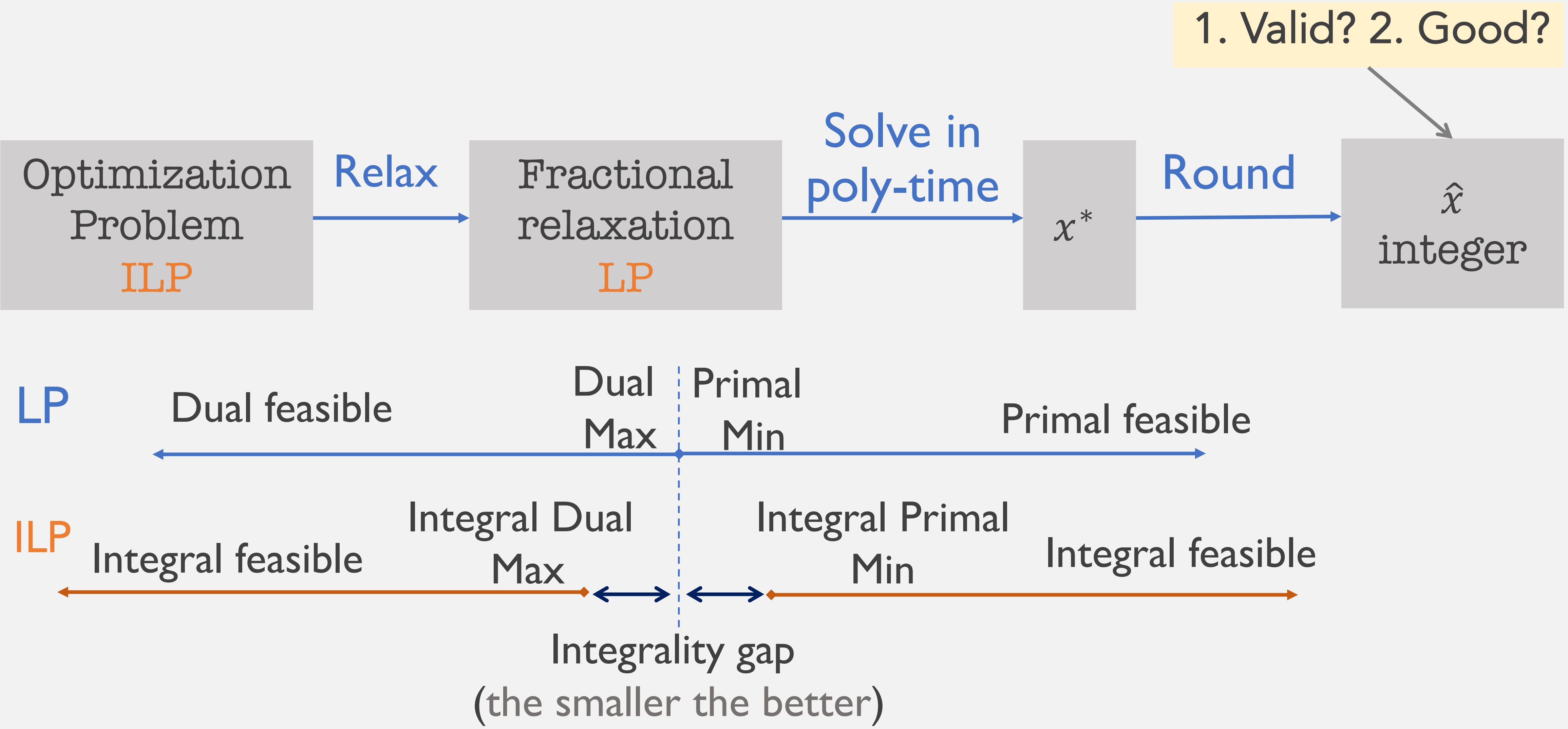
Let x^* be an optimal soln. for LP Σ
& optimal value $\mathrm{OPT} = \sum_i x_i^*$

⬅

$$x_i := \lfloor x_i^* \rceil = \begin{cases} 1, & \text{if } x_i^* \geq \frac{1}{2} \\ 0, & \text{otherwise} \end{cases}$$

- (Threshold) Rounding:
  i. $\{x_i\}$ is a feasible integral solution: $\forall (i,j) \in E, x_i^* \geq \frac{1}{2}$ or $x_j^* \geq \frac{1}{2}$ or both
  ii. $\sum_i x_i \leq \sum_i 2 \cdot x_i^* = 2 \cdot \mathrm{OPT} \leq 2 \cdot \mathrm{OPT}_{\mathrm{Int}}$ [optimal value of ILP Π, i.e. size of min vertex cover]

LP relaxation



Approximating set cover

Input. Set U of n elements, $S_1, \dots, S_m$ of subsets of U

Goal. Find $I \subseteq \{1, \dots, m\}$ of minimum size such that $\bigcup_{i \in I} S_i = U$

(ILP Π for Set cover)

For each $i \in \{1, \dots, m\}$, introduce $x_i \in \{0,1\}$

Min $\sum_{i=1}^{m} x_i$

Subject to:

$$\sum_{i: u \in S_i} x_i \geq 1, \qquad \forall u \in U$$

LP relaxation for set cover

(Set cover ILP Π)
Min $\sum_{i=1}^{m} x_i$
Subject to:

$$\sum_{i:u\in S_i} x_i \geq 1, \qquad \forall u \in U$$

$$x_i \in \{0,1\}, \qquad \forall i \in \{1,\dots,m\}$$

(Set cover Σ)
Min $\sum_{i=1}^{m} x_i$
Subject to:

$$\sum_{i:u\in S_i} x_i \geq 1, \qquad \forall u \in U$$

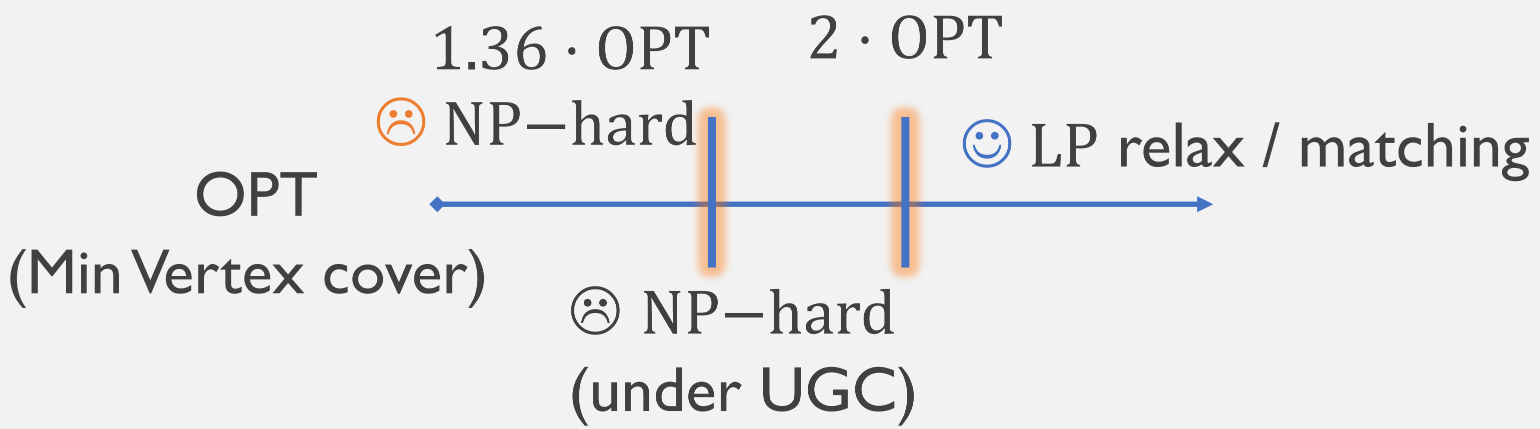
$$0 \leq x_i \leq 1, \forall i \in \{1,\dots,m\}$$

Let x^* be an optimal soln. for LP Σ & optimal value $\mathrm{OPT} = \sum_i x_i^*$

? $x_i := \lfloor x_i^* \rceil$

- Threshold rounding: does it cover all elements?
  - Ex. $u \in S_1, \dots, S_{100}; x_1^*, \dots x_{100}^* = \frac{1}{100} \Rightarrow x_1 = \cdots = x_{100} = 0$. u is missed!
- Randomized rounding! [Stay tuned]

Hardness of approximation



Theorem. It is NP-Hard to approximate Vertex Cover to with any factor below 1.36067. [i.e., otherwise, you can solve 3-SAT in poly-time]

Theorem'. It is NP-Hard to approximate Vertex Cover to with any factor below 2, assuming the unique games conjecture (UGC).

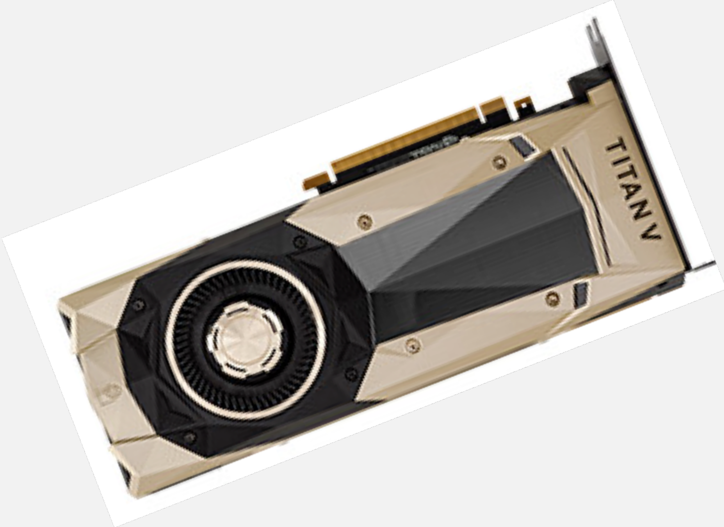
Want to read more?

https://cs.nyu.edu/~khot/papers/UGCSurvey.pdf

https://cs.stanford.edu/people/trevisan/pubs/inapprox.pdf

Scarce computational resources, which to invest on?

www.flickr.com



www.nvidia.com

www.computerhope.com

How about … coins?

Theorem. Randomness is useful

- Randomization. Allow fair coin flip in unit time

Power of randomness: primality testing

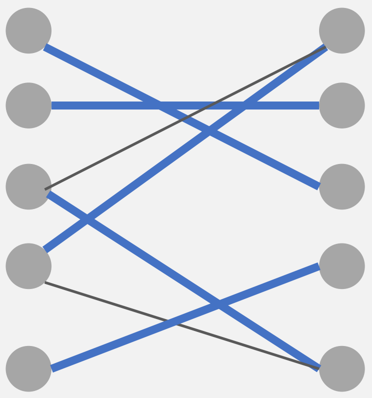
Is integer n Prime?

20,988,936,657,440,586,486,151,264,256,610,222,593,863,921

- Naive method: $O(n)$
- Randomized algorithm: Miller-Rabin 1977 $O(\log^4 n)$
- Deterministic algorithm: AKS 2002 $O(\log^{12} n)$

Miller-Rabin is still the way to go in practice!

Power of randomness: perfect matching



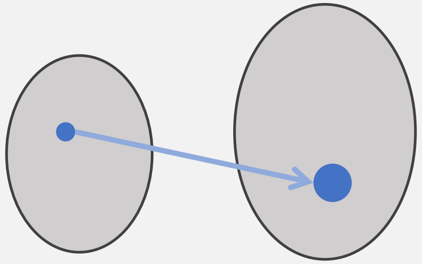
m: # edges
n: # nodes

- Deterministic algorithm: $O(nm)$
- Randomized algorithm: $O(\log^c nm)$

Exponentially faster!

Power of randomness beyond algorithm design

Probabilistic constructions



Nice error-correction codes exist:
random codes

Cryptography

Probabilistic Encryption*

Shafi Goldwasser and Silvio Micali

Probability 101

- (Discrete) Sample space $\Omega = \{\omega\}$
  - set of all possible outcomes of a random experiment
  - Event $E \subseteq \Omega$: a subset of the sample space
- Axioms of probability: a probability distribution is a mapping from events to real numbers $\Pr(\cdot): \mathcal{P}(\Omega) \to [0,1]$, satisfying
  - Probability of an event $\Pr(E) \geq 0$ for any event E
  - $\Pr(\Omega) = 1$
  - $\Pr(E \cup F) = \Pr(E) + \Pr(F)$ if $E \cap F = \emptyset$ (mutually exclusive)
- Ex. Roll a fair dice
  - $\Omega = \{1,2,3,4,5,6\}, \Pr(\omega) = \frac{1}{6}, \omega = 1, \dots, 6.$
  - $E = \{1,3,5\}$ dice being odd, & $\Pr(E) = 1/2$

N.B. $\bar{E} \coloneqq \Omega \backslash E$ complement event
$\Pr(\bar{E}) = 1 - \Pr(E)$

Probability 101 cont'd

- **Conditional probability:** $\Pr(B|A) := \frac{\Pr(A \cap B)}{\Pr(A)}$, assuming $\Pr(A) > 0$.

Bayes' theorem

Let E, F be two events and $\Pr(F) > 0$.

Then $\Pr(E|F) = \Pr(F|E) \cdot \frac{\Pr(E)}{\Pr(F)}$.

- **Independence:** Events A, B are independent iff. $\Pr(B|A) = \Pr(B)$.

i.e. $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$

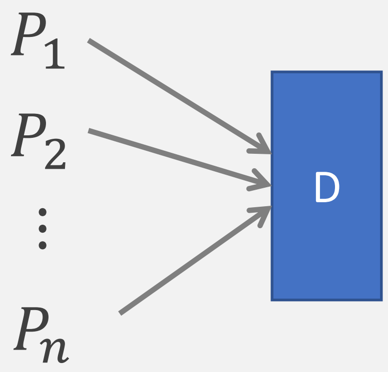
Contention resolution in a distributed system

Given: processes $P_1, \ldots, P_n$,

- each process competes for access to a shared database.
- If ≥ 2 processes access the database simultaneously, all processes are locked out.

Goal: a protocol so all processes get through on a regular basis

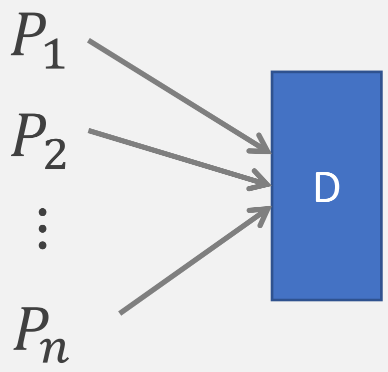
- Restriction: Processes can't communicate.



Contention resolution: randomized protocol

Protocol. Each process requests access to the database in round t with probability $p = 1/n$.

Theorem. All processes will succeed in accessing the database *at least once* within $O(n \ln n)$ rounds except with probability $\leq \frac{1}{n}$.



Randomized contention resolution: analysis 1

Def. $S[i,t]$ = event that process i succeeds in accessing the database in round t.

- Claim1. $\frac{1}{e \cdot n} \leq \Pr(S[i,t]) \leq \frac{1}{2n}$
- Pf. $\Pr(S[i,t]) = p(1-p)^{n-1}$

[Geometric distribution: independent Bernoulli trials]

Process i requests access (→ p)   None of remaining request access (→ $(1-p)^{n-1}$)

$$\Rightarrow \Pr(S[i,t]) = \frac{1}{n}(1-1/n)^{n-1} \in [\frac{1}{en}, \frac{1}{2n}] \quad [p = 1/n]$$

- $(1-1/n)^n$ converges monotonically from $1/4$ up to $1/e$.
- $(1-1/n)^{n-1}$ converges monotonically from $1/2$ down to $1/e$.

Randomized contention resolution: analysis 2

- Claim2. The probability that process i fails to access the database in $e \cdot n$ rounds is at most $1/e$. After $e \cdot n\ (c \ln n)$ rounds, the probability $\leq\ n^{-c}$.

- Pf. Let $F[i,t]$ = event that process i fails to access database in rounds 1 through t.

$$\Pr(F[i,t]) = \Pr\left(\overline{S[i,1]}\right) \cdot \ldots \cdot \Pr\left(\overline{S[i,t]}\right) \leq \left(1 - \frac{1}{en}\right)^t$$

[Independence & Claim 1]

- Choose $t = en$: $\Pr(F[i,t]) \leq \left(1 - \frac{1}{en}\right)^{en} \leq \frac{1}{e}$
- Choose $t = en \cdot c\ln n$: $\Pr(F[i,t]) \leq \left(\frac{1}{e}\right)^{c\ln n} \leq n^{-c}$

Randomized contention resolution: analysis 3

Theorem. All processes will succeed in accessing the database *at least once* within $2en \ln n$ rounds except with probability $\leq \frac{1}{n}$.

- Pf. Let $F[t]$ = event that some process fails to access database in rounds 1 through t.

Union Bound

Let E, F be two events. Then $\Pr(E \cup F) \leq \Pr(E) + \Pr(F)$.

$$\Pr(F[t]) = \Pr(\cup_{i=1}^{n} F[i,t]) \leq \sum_{i=1}^{n} \Pr(F[i,t]) \leq n \cdot \Pr(F[1,t])$$

- Choose $t = en \cdot 2\ln n$: $\Pr(F[t]) \leq n \cdot n^{-2} = 1/n$