CS 473: Algorithms, Fall 2018

Approximation Algorithms III

Lecture 10 September 24, 2018

10.1: Subset Sum

Subset Sum

Instance: $X = \{x_1, \ldots, x_n\} - n$ integer positive numbers, t - target number **Question**: \exists subset of X s.t. sum of its elements is t?

Assume x_1, \ldots, x_n are all $\leq n$. Then this problem can be solved in (A) The problem is still **NP-Hard**, so probably

exponential time.

(B) $O(n^3)$.

(C)
$$2^{O(\log^2 n)}$$

 $(\mathbf{D}) \quad O(m \log m)$

Subset Sum

Instance: $X = \{x_1, \ldots, x_n\} - n$ integer positive numbers, t - target number **Question**: \exists subset of X s.t. sum of its elements is t?

SolveSubsetSum (X, t, M) $b[0 \dots Mn] \leftarrow false$ // b[x] is true if x can be M: Max // realized by subset of X. value input $b[0] \leftarrow$ true. numbers. for $i = 1, \dots, n$ do for i = Mn down to x_i do

Subset Sum

Instance: $X = \{x_1, \ldots, x_n\} - n$ integer positive numbers, t - target number **Question**: \exists subset of X s.t. sum of its elements is t?

Subset Sum

Instance: $X = \{x_1, \ldots, x_n\} - n$ integer positive numbers, t - target number **Question**: \exists subset of X s.t. sum of its elements is t?

 $M: \qquad Max \qquad \begin{array}{c} \text{SolveSubsetSum } (X, t, M) \\ b[0 \dots Mn] \Leftarrow \text{false} \\ // b[x] \text{ is true if } x \text{ can be} \\ \text{Max} \qquad // \text{ realized by subset of } X. \\ \text{value input } b[0] \leftarrow \text{true.} \\ \text{for } i = 1, \dots, n \text{ do} \\ \text{for } i = Mn \text{ down to } x_i \text{ do} \end{array}$

Subset Sum

Instance: $X = \{x_1, \ldots, x_n\} - n$ integer positive numbers, t - target number **Question**: \exists subset of X s.t. sum of its elements is t?

 $M: \qquad Max \qquad \begin{array}{l} \text{SolveSubsetSum } (X, t, M) \\ Max \qquad \begin{array}{l} b[0 \dots Mn] \Leftarrow \text{false} \\ \text{// } b[x] \text{ is true if } x \text{ can be} \\ \text{numbers.} \\ \text{R.T.} \\ \text{O(M-2)} \end{array} \qquad \begin{array}{l} b[0] \leftarrow \text{true.} \\ \text{for } i = 1, \dots, n \text{ do} \\ \text{for } i = Mn \text{ down to } x_i \text{ do} \end{array}$

Efficient algorithm???

- 1. Algorithm solving **Subset Sum** in $O(Mn^2)$.
- 2. \boldsymbol{M} might be prohibitly large...
- 3. if $M = 2^n \implies$ algorithm is not polynomial time.
- 4. Subset Sum is NPC.
- 5. Still want to solve quickly even if M huge.
- 6. Optimization version:

Subset Sum Optimization

Instance: (X, t): A set X of n positive integers, and a target number t. **Question**: The largest number γ_{opt} one can represent as a subset sum of X which is smaller

Efficient algorithm???

- 1. Algorithm solving **Subset Sum** in $O(Mn^2)$.
- 2. \boldsymbol{M} might be prohibitly large...
- 3. if $M = 2^n \implies$ algorithm is not polynomial time.
- 4. Subset Sum is NPC.
- 5. Still want to solve quickly even if M huge.
- 6. Optimization version:

Subset Sum Optimization

Instance: (X, t): A set X of n positive integers, and a target number t. **Question**: The largest number γ_{opt} one can represent as a subset sum of X which is smaller

Efficient algorithm???

- 1. Algorithm solving **Subset Sum** in $O(Mn^2)$.
- 2. \boldsymbol{M} might be prohibitly large...
- 3. if $M = 2^n \implies$ algorithm is not polynomial time.
- 4. Subset Sum is NPC.
- 5. Still want to solve quickly even if M huge.
- 6. Optimization version:

Subset Sum Optimization

Instance: (X, t): A set X of n positive integers, and a target number t. **Question**: The largest number γ_{opt} one can represent as a subset sum of X which is smaller

2-approximation

Lemma

- 1. (X, t); Given instance of Subset Sum. $\gamma_{opt} \leq t$: Opt.
- 2. \implies Compute legal subset with sum $\geq \gamma_{\rm opt}/2$.
- 3. Running time $O(n \log n)$.

- 1. Sort numbers in X in decreasing order.
- 2. Greedily add numbers from largest to smallest (if possible).
- 3. *s*: Generates sum.

$\mathbf{2}\text{-}\mathsf{approx}\mathsf{imation}$

Lemma

- 1. (X, t); Given instance of Subset Sum. $\gamma_{\text{opt}} \leq t$: Opt.
- 2. \implies Compute legal subset with sum $\geq \gamma_{\rm opt}/2$.
- 3. Running time $O(n \log n)$.

- 1. Sort numbers in \boldsymbol{X} in decreasing order.
- 2. Greedily add numbers from largest to smallest (if possible).
- 3. *s*: Generates sum.

2-approximation

Lemma

- 1. (X, t); Given instance of Subset Sum. $\gamma_{\text{opt}} \leq t$: Opt.
- 2. \implies Compute legal subset with sum $\geq \gamma_{\rm opt}/2$.
- 3. Running time $O(n \log n)$.

- 1. Sort numbers in \boldsymbol{X} in decreasing order.
- 2. Greedily add numbers from largest to smallest (if possible).
- 3. *s*: Generates sum.

$\mathbf{2}\text{-}\mathsf{approx}\mathsf{imation}$

Lemma

- 1. (X, t); Given instance of Subset Sum. $\gamma_{\text{opt}} \leq t$: Opt.
- 2. \implies Compute legal subset with sum $\geq \gamma_{\rm opt}/2$.
- 3. Running time $O(n \log n)$.

- 1. Sort numbers in \boldsymbol{X} in decreasing order.
- 2. Greedily add numbers from largest to smallest (if possible).
- 3. *s*: Generates sum.

$\mathbf{2}\text{-}\mathsf{approx}\mathsf{imation}$

Lemma

- 1. (X, t); Given instance of Subset Sum. $\gamma_{\text{opt}} \leq t$: Opt.
- 2. \implies Compute legal subset with sum $\geq \gamma_{\rm opt}/2$.
- 3. Running time $O(n \log n)$.

- 1. Sort numbers in \boldsymbol{X} in decreasing order.
- 2. Greedily add numbers from largest to smallest (if possible).
- 3. s: Generates sum.

10.1.1: On the complexity of ε -approximation algorithms

Definition (

PROB: Maximization problem.

 $\varepsilon > 0$: approximation parameter. $\mathcal{A}(I, \varepsilon)$ is a *polynomial time approximation scheme* (**PTAS**) for **PROB**:

- $1. \ \forall I: \ (1-\varepsilon) \left| \mathsf{opt}(I) \right| \leq \left| \mathcal{A}(I,\varepsilon) \right| \leq \left| \mathsf{opt}(I) \right|,$
- 2. |opt(I)|: opt price,
- 3. $|\mathcal{A}(I,\varepsilon)|$: price of solution of \mathcal{A} .
- 4. \mathcal{A} running time polynomial in n for fixed ε .

For minimization problem: $|\mathsf{opt}(I)| \leq |\mathcal{A}(I,\varepsilon)| \leq (1+\varepsilon)|\mathsf{opt}(I)|.$

Definition (

PROB: Maximization problem.

 $\varepsilon > 0$: approximation parameter. $\mathcal{A}(I, \varepsilon)$ is a *polynomial time approximation scheme* (**PTAS**) for **PROB**:

- 1. $\forall I: (1-\varepsilon) \left| \mathsf{opt}(I) \right| \leq \left| \mathcal{A}(I,\varepsilon) \right| \leq \left| \mathsf{opt}(I) \right|,$
- 2. |opt(I)|: opt price,
- 3. $|\mathcal{A}(I,\varepsilon)|$: price of solution of \mathcal{A} .

4. \mathcal{A} running time polynomial in n for fixed ε .

For minimization problem:

 $|\mathsf{opt}(I)| \leq |\mathcal{A}(I,\varepsilon)| \leq (1+\varepsilon)|\mathsf{opt}(I)|.$

1. Example: Approximation algorithm with running time $O(n^{1/\varepsilon})$ is a PTAS. Algorithm with running time $O(1/\varepsilon^n)$ is not.

2. Fully polynomial...

Definition

- 3. Example: PTAS with running time $O(n^{1/\epsilon})$ is not a FPTAS.
- 4. Example: PTAS with running time $O(n^2/\varepsilon^3)$ is a FPTAS.

- Example: Approximation algorithm with running time O(n^{1/ε}) is a PTAS. Algorithm with running time O(1/εⁿ) is not.
- 2. Fully polynomial...

Definition (

- 3. Example: PTAS with running time $O(n^{1/\epsilon})$ is not a FPTAS.
- 4. Example: PTAS with running time $O(n^2/\varepsilon^3)$ is a FPTAS.

- Example: Approximation algorithm with running time O(n^{1/ε}) is a PTAS. Algorithm with running time O(1/εⁿ) is not.
- 2. Fully polynomial...

Definition (

- 3. Example: PTAS with running time $O(n^{1/\varepsilon})$ is not a FPTAS.
- 4. Example: PTAS with running time $O(n^2/\varepsilon^3)$ is a FPTAS.

- Example: Approximation algorithm with running time O(n^{1/ε}) is a PTAS. Algorithm with running time O(1/εⁿ) is not.
- 2. Fully polynomial...

Definition (

- 3. Example: PTAS with running time $O(n^{1/\varepsilon})$ is not a FPTAS.
- 4. Example: PTAS with running time $O(n^2/\varepsilon^3)$ is a FPTAS.

Approximating Subset Sum

Subset Sum Approx

Instance: (X, t, ε) : A set X of n positive integers, a target number t, and parameter $\varepsilon > 0$. **Question**: A number z that one can represent as a subset sum of X, such that $(1 - \varepsilon)\gamma_{\text{opt}} \leq z \leq \gamma_{\text{opt}} \leq t$.

Approximating Subset Sum

Looking again at the exact algorithm

 $\begin{array}{l} \textbf{ExactSubsetSum}(\textbf{S}, \ t) \\ n \leftarrow |S| \\ P_0 \leftarrow \{0\} \\ \textbf{for } i = 1 \dots n \ \textbf{do} \\ P_i \leftarrow P_{i-1} \cup (P_{i-1} + x_i) \\ \text{Remove from } P_i \ \textbf{all elements} > t \\ \end{array}$

1.
$$S = \{a_1, ..., a_n\}$$

 $x + S = \{a_1 + x, a_2 + x, ..., a_n + x\}$
2. Lists might explode in size.

Trim the lists...

Definition

L': Inc. sorted list of num-For two positive real numbers z < y, the bers number y is a $Trim(L', \delta)$ δ -approximation to z if $L = \langle y_1 \dots y_m \rangle$ $rac{\dot{y}}{1+\delta} \leq z \leq y.$ $curr \leftarrow y_1$ $L_{out} \leftarrow \{y_1\}$ for $i = 2 \dots m$ do Observation if $y_i > curr \cdot (1 + \delta)$ if $x \in L'$ then there Append y_i to L_{out} exists a number $y \in L_{out}$ such that $curr \leftarrow y_i$ return L_{out} $|y| \leq x \leq y(1+\delta),$ where

Trim the lists...

ApproxSubsetSum(S, t) $// \ S = \{x_1, \dots, x_n\}$, $Trim(L', \delta)$ // $x_1 \leq x_2 \leq \ldots < x_n$ $L = \langle y_1 \dots y_m \rangle$ $n \leftarrow |S|, \ L_0 \leftarrow \{0\},$ $curr \leftarrow y_1$ $L_{out} \leftarrow \{y_1\}$ $\delta = \varepsilon/2n$ for $i=2\ldots m$ do | for $i=1\ldots n$ do if $y_i > curr \cdot (1 + \delta) \mid E_i \leftarrow L_{i-1} \cup (L_{i-1} + x_i)$ Append y_i to $I_{out} \mid L_i \leftarrow \mathsf{Trim}(E_i, \delta)$ Remove from L_i elems > t. $curr \leftarrow y_i$ return L_{out} \overline{r} eturn largest element in L_n

E_i: Computed by merging two sorted lists in linear time.









14/41
















14/41









 \mathbf{n} $\mathbf{0}$

 $\mathbf{0}$

- 1. Can assume that trimmed lists L_i are sorted...
- 2. Algorithm: $E_i \leftarrow L_{i-1} \cup (L_{i-1} + x_i)$
- 3. So, this is just copy, shift, and merge of two sorted lists.
- 4. ... resulting in a sorted lest.
- 5. takes linear time in size of lists.

- 1. Can assume that trimmed lists L_i are sorted...
- 2. Algorithm: $E_i \leftarrow L_{i-1} \cup (L_{i-1} + x_i)$
- 3. So, this is just copy, shift, and merge of two sorted lists.
- 4. ... resulting in a sorted lest.
- 5. takes linear time in size of lists.

- 1. Can assume that trimmed lists L_i are sorted...
- 2. Algorithm: $E_i \leftarrow L_{i-1} \cup (L_{i-1} + x_i)$
- 3. So, this is just copy, shift, and merge of two sorted lists.
- 4. ... resulting in a sorted lest.
- 5. takes linear time in size of lists.

- 1. Can assume that trimmed lists L_i are sorted...
- 2. Algorithm: $E_i \leftarrow L_{i-1} \cup (L_{i-1} + x_i)$
- 3. So, this is just copy, shift, and merge of two sorted lists.
- 4. ... resulting in a sorted lest.
- 5. takes linear time in size of lists.

- 1. Can assume that trimmed lists L_i are sorted...
- 2. Algorithm: $E_i \leftarrow L_{i-1} \cup (L_{i-1} + x_i)$
- 3. So, this is just copy, shift, and merge of two sorted lists.
- 4. ... resulting in a sorted lest.
- 5. takes linear time in size of lists.

Analysis

- 1. E_i list generated by algorithm in *i*th iteration.
- 2. P_i : list of numbers (no trimming).

Claim

For any $x \in P_i$ there exists $y \in L_i$ such that $y \leq x \leq (1+\delta)^i y$. Proof

- 1. If $x \in P_1$ then follows by observation above.
- 2. If $x \in P_{i-1} \Longrightarrow$ (induction) $\exists y' \in L_{i-1}$ s.t. $y' \leq x \leq (1+\delta)^{i-1}y'$.
- 3. By observation $\exists y \in L_i$ s.t. $y \leq y' \leq (1+\delta)y$, As such,

Analysis

- 1. E_i list generated by algorithm in *i*th iteration.
- 2. P_i : list of numbers (no trimming).

Claim

For any $x \in P_i$ there exists $y \in L_i$ such that $y \leq x \leq (1 + \delta)^i y$. Proof

- 1. If $x \in P_1$ then follows by observation above.
- 2. If $x \in P_{i-1} \implies$ (induction) $\exists y' \in L_{i-1}$ s.t. $y' \leq x \leq (1+\delta)^{i-1}y'$.
- 3. By observation $\exists y \in L_i$ s.t. $y \leq y' \leq (1+\delta)y$, As such,

Analysis

- 1. E_i list generated by algorithm in *i*th iteration.
- 2. P_i : list of numbers (no trimming).

Claim

For any $x \in P_i$ there exists $y \in L_i$ such that $y \leq x \leq (1 + \delta)^i y$. Proof

- 1. If $x \in P_1$ then follows by observation above.
- 2. If $x \in P_{i-1} \implies$ (induction) $\exists y' \in L_{i-1}$ s.t. $y' \leq x \leq (1+\delta)^{i-1}y'$.
- 3. By observation $\exists y \in L_i$ s.t. $y \leq y' \leq (1 + \delta)y$, As such,

Proof continued

- 1. If $x \in P_i \setminus P_{i-1} \implies x = \alpha + x_i$, for some $\alpha \in P_{i-1}$.
- 2. By induction, $\exists \alpha' \in L_{i-1}$ s.t. $\alpha' \leq \alpha \leq (1+\delta)^{i-1} \alpha'.$
- 3. Thus, $lpha' + x_i \in E_i$.
- 4. $\exists x' \in L_i ext{ s.t. } x' \leq lpha' + x_i \leq (1+\delta)x'.$
- 5. Thus,

 $\begin{array}{l} x'\leq \alpha'+x_i\leq \alpha+x_i=x\leq (1\!+\!\delta)^{i-1}\alpha'+x_i\leq \\ (1\!+\!\delta)^{i-1}(\alpha'+x_i)\leq (1\!+\!\delta)^ix'. \end{array}$

Proof continued

- 1. If $x \in P_i \setminus P_{i-1} \implies x = \alpha + x_i$, for some $\alpha \in P_{i-1}$.
- 2. By induction, $\exists \alpha' \in L_{i-1}$ s.t. $\alpha' \leq \alpha \leq (1 + \delta)^{i-1} \alpha'.$
- 3. Thus, $\alpha' + x_i \in E_i$.
- 4. $\exists x' \in L_i$ s.t. $x' \leq lpha' + x_i \leq (1+\delta)x'$.
- 5. Thus,

 $\begin{array}{l} x' \leq \alpha' + x_i \leq \alpha + x_i = x \leq (1 + \delta)^{i - 1} \alpha' + x_i \leq \\ (1 + \delta)^{i - 1} (\alpha' + x_i) \leq (1 + \delta)^i x'. \end{array}$

Proof continued

- 1. If $x \in P_i \setminus P_{i-1} \implies x = \alpha + x_i$, for some $\alpha \in P_{i-1}$.
- 2. By induction, $\exists \alpha' \in L_{i-1}$ s.t. $\alpha' \leq \alpha \leq (1 + \delta)^{i-1} \alpha'.$
- 3. Thus, $\alpha' + x_i \in E_i$.
- 4. $\exists x' \in L_i ext{ s.t. } x' \leq lpha' + x_i \leq (1+\delta)x'.$
- 5. Thus,

 $egin{aligned} x' &\leq lpha' + x_i \leq lpha + x_i = x \leq (1 + \delta)^{i - 1} lpha' + x_i \leq (1 + \delta)^{i - 1} (lpha' + x_i) \leq (1 + \delta)^i x'. \end{aligned}$

Proof continued

- 1. If $x \in P_i \setminus P_{i-1} \implies x = \alpha + x_i$, for some $\alpha \in P_{i-1}$.
- 2. By induction, $\exists \alpha' \in L_{i-1}$ s.t. $\alpha' \leq \alpha \leq (1 + \delta)^{i-1} \alpha'.$
- 3. Thus, $\alpha' + x_i \in E_i$.
- 4. $\exists x' \in L_i ext{ s.t. } x' \leq lpha' + x_i \leq (1+\delta)x'.$
- 5. Thus,

 $egin{aligned} & m{x'} \leq m{lpha'} + m{x_i} \leq m{lpha} + m{x_i} = x \leq (1 + \delta)^{i - 1} m{lpha'} + m{x_i} \leq (1 + \delta)^{i - 1} m{lpha'} + m{x_i}) \leq (1 + \delta)^i m{x'}. \end{aligned}$

Proof continued

- 1. If $x \in P_i \setminus P_{i-1} \implies x = \alpha + x_i$, for some $\alpha \in P_{i-1}$.
- 2. By induction, $\exists \alpha' \in L_{i-1}$ s.t. $\alpha' \leq \alpha \leq (1 + \delta)^{i-1} \alpha'.$
- 3. Thus, $\alpha' + x_i \in E_i$.
- 4. $\exists x' \in L_i ext{ s.t. } x' \leq lpha' + x_i \leq (1+\delta)x'.$
- 5. Thus,

 $egin{aligned} & x'\leq lpha'+x_i\leq lpha+x_i=x\leq (1\!+\!\delta)^{i-1}lpha'+x_i\leq \ & (1+\delta)^{i-1}(lpha'+x_i)\leq (1+\delta)^ix'. \end{aligned}$

Proof continued

- 1. If $x \in P_i \setminus P_{i-1} \implies x = \alpha + x_i$, for some $\alpha \in P_{i-1}$.
- 2. By induction, $\exists \alpha' \in L_{i-1}$ s.t. $\alpha' \leq \alpha \leq (1 + \delta)^{i-1} \alpha'.$
- 3. Thus, $\alpha' + x_i \in E_i$.
- 4. $\exists x' \in L_i ext{ s.t. } x' \leq lpha' + x_i \leq (1+\delta)x'.$
- 5. Thus,

 $egin{aligned} x' &\leq lpha' + x_i \leq lpha + x_i = x \leq (1 + \delta)^{i - 1} lpha' + x_i \leq \ (1 + \delta)^{i - 1} (lpha' + x_i) \leq (1 + \delta)^i x'. \end{aligned}$

Proof continued

- 1. If $x \in P_i \setminus P_{i-1} \implies x = \alpha + x_i$, for some $\alpha \in P_{i-1}$.
- 2. By induction, $\exists \alpha' \in L_{i-1}$ s.t. $\alpha' \leq \alpha \leq (1 + \delta)^{i-1} \alpha'.$
- 3. Thus, $\alpha' + x_i \in E_i$.
- 4. $\exists x' \in L_i ext{ s.t. } x' \leq lpha' + x_i \leq (1+\delta)x'.$
- 5. Thus,

 $x' \leq lpha' + x_i \leq lpha + x_i = x \leq (1+\delta)^{i-1} lpha' + x_i \leq (1+\delta)^{i-1} (lpha' + x_i) \leq (1+\delta)^i x'.$

Proof continued

- 1. If $x \in P_i \setminus P_{i-1} \implies x = \alpha + x_i$, for some $\alpha \in P_{i-1}$.
- 2. By induction, $\exists \alpha' \in L_{i-1}$ s.t. $\alpha' \leq \alpha \leq (1 + \delta)^{i-1} \alpha'.$
- 3. Thus, $\alpha' + x_i \in E_i$.
- 4. $\exists x' \in L_i ext{ s.t. } x' \leq lpha' + x_i \leq (1+\delta)x'.$
- 5. Thus,

 $\begin{array}{l} x' \leq \alpha' + x_i \leq \alpha + x_i = x \leq (1 + \delta)^{i - 1} \alpha' + x_i \leq \\ (1 + \delta)^{i - 1} (\alpha' + x_i) \leq (1 + \delta)^i x'. \end{array}$

$10.1.1.1: {\sf Running time}$

Running time of ApproxSubsetSum

Lemma For $x \in [0, 1]$, it holds $\exp(x/2) \le (1 + x)$. Lemma For $0 < \delta < 1$, and $x \ge 1$, we have $\log_{1+\delta} x \le \frac{2\ln x}{\delta} = O\left(\frac{\ln x}{\delta}\right)$.

See notes for a proof of lemmas.

Running time of ApproxSubsetSum

Observation

In a list generated by Trim, for any number x, there are no two numbers in the trimmed list between x and $(1 + \delta)x$.

Lemma $|L_i| = O\Bigl((n/arepsilon) \log n\Bigr)$, for $i=1,\ldots,n$.

Running time of ApproxSubsetSum Proof.

- 1. $L_{i-1} + x_i \subseteq [x_i, ix_i].$
- 2. Trimming $L_{i-1} + x_i$ results in list of size

$$\log_{1+\delta}rac{ix_i}{x_i}=Oigg(rac{\ln i}{\delta}igg)=Oigg(rac{\ln n}{\delta}igg),$$

3. Now, $\delta = \varepsilon/2n$, and

$$egin{aligned} |L_i| &\leq |L_{i-1}| + Oigg(rac{\ln n}{\delta}igg) \leq |L_{i-1}| + Oigg(rac{n\ln n}{arepsilon}igg) \ &= Oigg(rac{n^2\log n}{arepsilon}igg) \end{aligned}$$

Running time of ApproxSubsetSum Proof.

- 1. $L_{i-1} + x_i \subseteq [x_i, ix_i]$.
- 2. Trimming $L_{i-1} + x_i$ results in list of size

$$\log_{1+\delta}rac{ix_i}{x_i}=Oigg(rac{\ln i}{\delta}igg)=Oigg(rac{\ln n}{\delta}igg),$$

3. Now, $\delta = \varepsilon/2n$, and

$$egin{aligned} |L_i| &\leq |L_{i-1}| + Oigg(rac{\ln n}{\delta}igg) \leq |L_{i-1}| + Oigg(rac{n\ln n}{arepsilon}igg) \ &= Oigg(rac{n^2\log n}{arepsilon}igg) \ \end{aligned}$$

Running time of ApproxSubsetSum

Lemma The running time of ApproxSubsetSum is $O\left(\frac{n^3}{\varepsilon}\log n\right).$

Proof.

- 1. Running time of **ApproxSubsetSum** dominated by total length of L_1, \ldots, L_n .
- 2. Above lemma implies

$$\sum_i |L_i| = Oigg(n imes rac{n^2}{arepsilon} \log nigg) = Oigg(rac{n^3}{arepsilon} \log nigg)$$

3. Trim runs in time proportional to size of lists.

$$(0) | D T $O(n^3)$$$

ApproxSubsetSum

Theorem **ApproxSubsetSum** returns $u \leq t$, s.t. $\frac{\gamma_{\text{opt}}}{1+\varepsilon} \leq u \leq \gamma_{\text{opt}} \leq t$, γ_{opt} : opt solution. Running time is $O((n^3/\varepsilon) \log n)$.

Proof.

- 1. Running time from above.
- 2. $\gamma_{ ext{opt}} \in P_n$: optimal solution.
- 3. $\exists z \in L_n$, such that $z \leq \mathsf{opt} \leq (1+\delta)^n z$
- 4. $(1+\delta)^n = (1+\varepsilon/2n)^n \le \exp\left(\frac{\varepsilon}{2}\right) \le 1+\varepsilon,$ since $1+x \le e^x$ for $x \ge 0$.
- $5 \sim 1(1 \perp c) \leq \alpha \leq \text{opt} \leq t$
Theorem **ApproxSubsetSum** returns $u \leq t$, s.t. $\frac{\gamma_{\text{opt}}}{1+\varepsilon} \leq u \leq \gamma_{\text{opt}} \leq t$, γ_{opt} : opt solution. Running time is $O((n^3/\varepsilon) \log n)$.

Proof.

- 1. Running time from above.
- 2. $\gamma_{\text{opt}} \in P_n$: optimal solution.
- 3. $\exists z \in L_n$, such that $z \leq \mathsf{opt} \leq (1+\delta)^n z$
- 4. $(1+\delta)^n = (1+\varepsilon/2n)^n \le \exp\left(\frac{\varepsilon}{2}\right) \le 1+\varepsilon,$ since $1+x \le e^x$ for $x \ge 0$.
- $5 \sim 1/(1 + c) \leq \alpha \leq \text{ont} \leq t$

Theorem **ApproxSubsetSum** returns $u \leq t$, s.t. $\frac{\gamma_{\text{opt}}}{1+\varepsilon} \leq u \leq \gamma_{\text{opt}} \leq t$, γ_{opt} : opt solution. Running time is $O((n^3/\varepsilon) \log n)$.

Proof.

- 1. Running time from above.
- 2. $\gamma_{\text{opt}} \in P_n$: optimal solution.
- 3. $\exists z \in L_n$, such that $z \leq \text{opt} \leq (1+\delta)^n z$

4. $(1+\delta)^n = (1+\varepsilon/2n)^n \le \exp\left(\frac{\varepsilon}{2}\right) \le 1+\varepsilon,$ since $1+x \le e^x$ for $x \ge 0$.

5 $\alpha = 1/(1 \perp c) \leq \pi \leq \text{ont} \leq t$

Theorem **ApproxSubsetSum** returns $u \leq t$, s.t. $\frac{\gamma_{\text{opt}}}{1+\varepsilon} \leq u \leq \gamma_{\text{opt}} \leq t$, γ_{opt} : opt solution. Running time is $O((n^3/\varepsilon) \log n)$.

Proof.

- 1. Running time from above.
- 2. $\gamma_{\text{opt}} \in P_n$: optimal solution.
- 3. $\exists z \in L_n$, such that $z \leq \mathsf{opt} \leq (1+\delta)^n z$
- 4. $(1+\delta)^n = (1+\varepsilon/2n)^n \le \exp\left(\frac{\varepsilon}{2}\right) \le 1+\varepsilon$, since $1+x \le e^x$ for $x \ge 0$.

 $5 \propto 1/(1 \perp \epsilon) \leq \pi \leq \text{ont} \leq t$

Theorem **ApproxSubsetSum** returns $u \leq t$, s.t. $\frac{\gamma_{\text{opt}}}{1+\varepsilon} \leq u \leq \gamma_{\text{opt}} \leq t$, γ_{opt} : opt solution. Running time is $O((n^3/\varepsilon) \log n)$.

Proof.

- 1. Running time from above.
- 2. $\gamma_{\text{opt}} \in P_n$: optimal solution.
- 3. $\exists z \in L_n$, such that $z \leq \text{opt} \leq (1+\delta)^n z$
- 4. $(1+\delta)^n = (1+\varepsilon/2n)^n \le \exp\left(\frac{\varepsilon}{2}\right) \le 1+\varepsilon,$ since $1+x \le e^x$ for $x \ge 0$.
- $5 \sim ((1 + c) \leq \pi \leq \text{ont} \leq t)$

10.2: Maximal matching

Maximal matching

- 1. $\mathbf{G} = (\mathbf{V}, \mathbf{E})$
- 2. Compute maximal matching...
- 3. $X \subseteq \mathbf{E}$ which is maximal and independent.
- 4. Maximal = can not improved by adding an edge.
- 5. Maximum = largest possible set among all possible sets.
- 6. Computing the maximum is hard then computing maximal solution.
- 7. Q: Find maximal matching quickly and of large size...

































6



- 1. Algorithm: Repeatedly pick an arbitrary edge and remove it.
- 2. M: Generated matching. X: Maximal matching.
- 3. Clearly a maximal matching...
- 4. This is a **2**-approximation to the maximum matching.
- 5. Because...
- 6. Every edge in M "kills" two edges of X in the worst case.

- 1. Algorithm: Repeatedly pick an arbitrary edge and remove it.
- 2. M: Generated matching. X: Maximal matching.
- 3. Clearly a maximal matching...
- 4. This is a 2-approximation to the maximum matching.
- 5. Because...
- 6. Every edge in M "kills" two edges of X in the worst case.

- 1. Algorithm: Repeatedly pick an arbitrary edge and remove it.
- 2. M: Generated matching. X: Maximal matching.
- 3. Clearly a maximal matching...
- 4. This is a 2-approximation to the maximum matching.
- 5. Because...
- 6. Every edge in M "kills" two edges of X in the worst case.

- 1. Algorithm: Repeatedly pick an arbitrary edge and remove it.
- 2. M: Generated matching. X: Maximal matching.
- 3. Clearly a maximal matching...
- 4. This is a 2-approximation to the maximum matching.
- 5. Because...
- 6. Every edge in M "kills" two edges of X in the worst case.

- 1. Algorithm: Repeatedly pick an arbitrary edge and remove it.
- 2. M: Generated matching. X: Maximal matching.
- 3. Clearly a maximal matching...
- 4. This is a 2-approximation to the maximum matching.
- 5. Because...
- 6. Every edge in M "kills" two edges of X in the worst case.

Maximal matching: Result

Theorem

Given a graph **G** one can compute in O(n + m) time, a maximal matching with at least |X|/2 edges, where X is the size of the maximum (optimal) matching.

10.2.1: Bin packing

Bin packing

Problem definition

Bin Packing

Instance: v: Bin size. $S = \{\alpha_1, \dots, \alpha_n\}$: nitems α_i : size of *i*th item. **Target**: Find min # B, and a decomposition S_1, \dots, S_B of S, such that $\forall j$ $\sum_{x \in S_i} \leq v$.

- 1. $\cup_i S_i = S$ and $\forall i \neq j \quad S_i \cap S_j = \emptyset$.
- 2. NP-Hard from Partition.
- 3. **NP-Hard** to approximate within 3/2.
- 4. Natural problem...
- 5. How to approximate?

Bin packing

Problem definition

Bin Packing

Instance: *v*: Bin size. $S = \{\alpha_1, \dots, \alpha_n\}$: *n* items α_i : size of *i*th item. **Target**: Find min # B, and a decomposition S_1, \dots, S_B of *S*, such that $\forall j$ $\sum_{x \in S_i} \leq v$.

- 1. $\cup_i S_i = S$ and $\forall i \neq j \quad S_i \cap S_j = \emptyset$. 2. NP-Hard from Partition
- 3. **NP-Hard** to approximate within 3/2.
- 4. Natural problem...
- 5. How to approximate?

Bin packing

Problem definition

Bin Packing

Instance: *v*: Bin size. $S = \{\alpha_1, \dots, \alpha_n\}$: *n* items α_i : size of *i*th item. **Target**: Find min # B, and a decomposition S_1, \dots, S_B of *S*, such that $\forall j$ $\sum_{x \in S_i} \leq v$.

- 1. $\cup_i S_i = S$ and $\forall i \neq j \quad S_i \cap S_j = \emptyset$.
- 2. NP-Hard from Partition.
- 3. **NP-Hard** to approximate within 3/2.
- 4. Natural problem...
- 5. How to approximate?
Problem definition

Bin Packing

Instance: *v*: Bin size. $S = \{\alpha_1, \dots, \alpha_n\}$: *n* items α_i : size of *i*th item. **Target**: Find min # B, and a decomposition S_1, \dots, S_B of *S*, such that $\forall j$ $\sum_{x \in S_i} \leq v$.

- 1. $\cup_i S_i = S$ and $\forall i \neq j \quad S_i \cap S_j = \emptyset$.
- 2. NP-Hard from Partition.
- 3. **NP-Hard** to approximate within 3/2.
- 4. Natural problem...
- 5. How to approximate?

Problem definition

Bin Packing

Instance: *v*: Bin size. $S = \{\alpha_1, \dots, \alpha_n\}$: *n* items α_i : size of *i*th item. **Target**: Find min # B, and a decomposition S_1, \dots, S_B of *S*, such that $\forall j$ $\sum_{x \in S_i} \leq v$.

1.
$$\cup_i S_i = S$$
 and $\forall i \neq j \quad S_i \cap S_j = \emptyset$.

- 2. NP-Hard from Partition.
- 3. **NP-Hard** to approximate within 3/2.
- 4. Natural problem...
- 5. How to approximate?

Problem definition

Bin Packing

Instance: *v*: Bin size. $S = \{\alpha_1, \dots, \alpha_n\}$: *n* items α_i : size of *i*th item. **Target**: Find min # B, and a decomposition S_1, \dots, S_B of *S*, such that $\forall j$ $\sum_{x \in S_i} \leq v$.

1.
$$\cup_i S_i = S$$
 and $\forall i \neq j \quad S_i \cap S_j = \emptyset$.

- 2. NP-Hard from Partition.
- 3. **NP-Hard** to approximate within 3/2.
- 4. Natural problem...
- 5. How to approximate?

Problem definition

Bin Packing

Instance: *v*: Bin size. $S = \{\alpha_1, \ldots, \alpha_n\}$: *n* items α_i : size of *i*th item. **Target**: Find min # B, and a decomposition S_1, \ldots, S_B of *S*, such that $\forall j \qquad \sum_{x \in S_i} \leq v$.

- 1. $\cup_i S_i = S$ and $\forall i \neq j \quad S_i \cap S_j = \emptyset$.
- 2. NP-Hard from Partition.
- 3. **NP-Hard** to approximate within 3/2.
- 4. Natural problem...
- 5. How to approximate?

Problem definition

Bin Packing

Instance: *v*: Bin size. $S = \{\alpha_1, \ldots, \alpha_n\}$: *n* items α_i : size of *i*th item. **Target**: Find min # B, and a decomposition S_1, \ldots, S_B of *S*, such that $\forall j \qquad \sum_{x \in S_i} \leq v$.

- 1. $\cup_i S_i = S$ and $\forall i \neq j \quad S_i \cap S_j = \emptyset$.
- 2. NP-Hard from Partition.
- 3. **NP-Hard** to approximate within 3/2.
- 4. Natural problem...
- 5. How to approximate?

Bin packing: First fit Analysis

Lemma First fit is a 2-approximation.

Proof.

Observe that only one bin can have less than v/2 content in it...



10.3: Independent set of axis-parallel rectangles





An example



Assume: Open rectangles.



Independent set of rectangles.

Independent set of intervals

Clicker question

Given n intervals on the real line, computing the largest independent set of intervals on the real line, can be done in:

- (A) O(n) time.
- (B) $O(n \log n)$ time.
- (C) $O(n^{3/2})$ time.
- (D) $O(n^2)$ time.
- (E) NP-Hard.

















Algorithm: Divide & Conquer

 \mathcal{R} : A set of axis parallel rectangles.



1. If $S_M > \operatorname{Opt}/(2 \lg n)$... done. 4. $S_L + S_R \ge \frac{(1-1/(2 \lg n)) \operatorname{Opt}}{2 \lg (n/2)}$ $\frac{2 \lg n - 2}{(2 \lg n)(2 \lg n - 2)} \ge \frac{1}{2 \lg n}$

1. If
$$S_M \ge \operatorname{Opt}/(2 \lg n)$$
... done.
2. $\operatorname{Opt}_L + \operatorname{Opt}_R \ge (1 - 1/(2 \lg n))\operatorname{Opt}$.
3. By induction: $S_L \ge \operatorname{Opt}_L/(2 \lg (n/2))$ and $S_R \ge \operatorname{Opt}_R/(2 \lg (n/2))$.
4. $S_L + S_R \ge \frac{(1 - 1/(2 \lg n))\operatorname{Opt}}{2 \lg (n/2)}$
5. $\frac{(1 - 1/(2 \lg n))}{2 \lg (n/2)} = \frac{1}{2 \lg n - 2} = \frac{1}{(2 \lg n)(2 \lg n - 2)}$
 $\ge \frac{2 \lg n - 1}{(2 \lg n)(2 \lg n - 2)} \ge \frac{2 \lg n - 2}{(2 \lg n)(2 \lg n - 2)} \ge \frac{1}{2 \lg n - 2}$

1. If
$$S_M \ge \operatorname{Opt}/(2 \lg n)$$
... done.
2. $\operatorname{Opt}_L + \operatorname{Opt}_R \ge (1 - 1/(2 \lg n))\operatorname{Opt}$.
3. By induction: $S_L \ge \operatorname{Opt}_L/(2 \lg (n/2))$ and $S_R \ge \operatorname{Opt}_R/(2 \lg (n/2))$.
4. $S_L + S_R \ge \frac{(1 - 1/(2 \lg n))\operatorname{Opt}}{2 \lg (n/2)}$
5. $\frac{(1 - 1/(2 \lg n))}{2 \lg (n/2)} = \frac{1}{2 \lg n - 2} - \frac{1}{(2 \lg n)(2 \lg n - 2)}$
 $\ge \frac{2 \lg n - 1}{(2 \lg n)(2 \lg n - 2)} \ge \frac{2 \lg n - 2}{(2 \lg n)(2 \lg n - 2)} \ge \frac{1}{2 \lg n - 2}$

1. If
$$S_M \ge \operatorname{Opt}/(2 \lg n)$$
... done.
2. $\operatorname{Opt}_L + \operatorname{Opt}_R \ge (1 - 1/(2 \lg n))\operatorname{Opt}$.
3. By induction: $S_L \ge \operatorname{Opt}_L/(2 \lg (n/2))$ and $S_R \ge \operatorname{Opt}_R/(2 \lg (n/2))$.
4. $S_L + S_R \ge \frac{(1 - 1/(2 \lg n))\operatorname{Opt}}{2 \lg (n/2)}$
5. $\frac{(1 - 1/(2 \lg n))}{2 \lg (n/2)} = \frac{1}{2 \lg n - 2} = \frac{2 \lg n - 1}{(2 \lg n)(2 \lg n - 2)} \ge \frac{2 \lg n - 2}{(2 \lg n)(2 \lg n - 2)} \ge \frac{2 \lg n - 2}{(2 \lg n)(2 \lg n - 2)} \ge \frac{1}{2 \lg n - 2}$

1. If
$$S_M \ge \operatorname{Opt}/(2 \lg n)$$
... done.
2. $\operatorname{Opt}_L + \operatorname{Opt}_R \ge (1 - 1/(2 \lg n))\operatorname{Opt}$.
3. By induction: $S_L \ge \operatorname{Opt}_L/(2 \lg (n/2))$ and $S_R \ge \operatorname{Opt}_R/(2 \lg (n/2))$.
4. $S_L + S_R \ge \frac{(1 - 1/(2 \lg n))\operatorname{Opt}}{2 \lg (n/2)}$
5. $\frac{(1 - 1/(2 \lg n))}{2 \lg (n/2)} = \frac{1}{2 \lg n - 2} - \frac{1}{(2 \lg n)(2 \lg n - 2)}$
 $\ge \frac{2 \lg n - 1}{(2 \lg n)(2 \lg n - 2)} \ge \frac{2 \lg n - 2}{(2 \lg n)(2 \lg n - 2)} \ge \frac{1}{2 \lg n - 2}$

1. If
$$S_M \ge \operatorname{Opt}/(2 \lg n)$$
... done.
2. $\operatorname{Opt}_L + \operatorname{Opt}_R \ge (1 - 1/(2 \lg n))\operatorname{Opt}$.
3. By induction: $S_L \ge \operatorname{Opt}_L/(2 \lg (n/2))$ and $S_R \ge \operatorname{Opt}_R/(2 \lg (n/2))$.
4. $S_L + S_R \ge \frac{(1 - 1/(2 \lg n))\operatorname{Opt}}{2 \lg (n/2)}$
5. $\frac{(1 - 1/(2 \lg n))}{2 \lg (n/2)} = \frac{1}{2 \lg n - 2} - \frac{1}{(2 \lg n)(2 \lg n - 2)}$
 $\ge \frac{2 \lg n - 1}{(2 \lg n)(2 \lg n - 2)} \ge \frac{2 \lg n - 2}{(2 \lg n)(2 \lg n - 2)} \ge \frac{1}{2 \lg n - 2}$

1. If
$$S_M \ge \operatorname{Opt}/(2 \lg n)$$
... done.
2. $\operatorname{Opt}_L + \operatorname{Opt}_R \ge (1 - 1/(2 \lg n))\operatorname{Opt}$.
3. By induction: $S_L \ge \operatorname{Opt}_L/(2 \lg (n/2))$ and $S_R \ge \operatorname{Opt}_R/(2 \lg (n/2))$.
4. $S_L + S_R \ge \frac{(1 - 1/(2 \lg n))\operatorname{Opt}}{2 \lg (n/2)}$
5. $\frac{(1 - 1/(2 \lg n))}{2 \lg (n/2)} = \frac{1}{2 \lg n - 2} - \frac{1}{(2 \lg n)(2 \lg n - 2)}$
 $\ge \frac{2 \lg n - 1}{(2 \lg n)(2 \lg n - 2)} \ge \frac{2 \lg n - 2}{(2 \lg n)(2 \lg n - 2)} \ge \frac{1}{2 \lg n - 2}$

1. If
$$S_M \ge \operatorname{Opt}/(2 \lg n)$$
... done.
2. $\operatorname{Opt}_L + \operatorname{Opt}_R \ge (1 - 1/(2 \lg n))\operatorname{Opt}$.
3. By induction: $S_L \ge \operatorname{Opt}_L/(2 \lg (n/2))$ and $S_R \ge \operatorname{Opt}_R/(2 \lg (n/2))$.
4. $S_L + S_R \ge \frac{(1 - 1/(2 \lg n))\operatorname{Opt}}{2 \lg (n/2)}$
5. $\frac{(1 - 1/(2 \lg n))}{2 \lg (n/2)} = \frac{1}{2 \lg n - 2} - \frac{1}{(2 \lg n)(2 \lg n - 2)}$
 $\ge \frac{2 \lg n - 1}{(2 \lg n)(2 \lg n - 2)} \ge \frac{2 \lg n - 2}{(2 \lg n)(2 \lg n - 2)} \ge \frac{1}{2 \lg n - 2}$

1. If
$$S_M \ge \operatorname{Opt}/(2 \lg n)$$
... done.
2. $\operatorname{Opt}_L + \operatorname{Opt}_R \ge (1 - 1/(2 \lg n))\operatorname{Opt}$.
3. By induction: $S_L \ge \operatorname{Opt}_L/(2 \lg (n/2))$ and $S_R \ge \operatorname{Opt}_R/(2 \lg (n/2))$.
4. $S_L + S_R \ge \frac{(1 - 1/(2 \lg n))\operatorname{Opt}}{2 \lg (n/2)}$
5. $\frac{(1 - 1/(2 \lg n))}{2 \lg (n/2)} = \frac{1}{2 \lg n - 2} - \frac{1}{(2 \lg n)(2 \lg n - 2)}$
 $\ge \frac{2 \lg n - 1}{(2 \lg n)(2 \lg n - 2)} \ge \frac{2 \lg n - 2}{(2 \lg n)(2 \lg n - 2)} \ge \frac{1}{2 \lg n - 2}$