Cormen Section 35.5

Recall Subset sum:

Given $S = \{x_1, x_2, \dots, x_n\}$ and $t \in \mathbb{Z}_+$

Question $\exists ?\ I \subseteq \{1, 2, \dots, n\}$ s.t $\sum_{i \in I} x_i = t$?

Optimization version: Find maximum $t^* \leq t$ s.t $\exists\, I \subseteq \{1, 2 \dots n\}$ with $\sum_{i \in I} x_i = t^*$

Definition: A polynomial approximation scheme for a problem Q is an algorithm $A = A(\varepsilon)$ such that A finds a solution which is with $(1+\varepsilon)$ of opt solution:

$$\max\left(\frac{C}{C^*}, \frac{C^*}{C}\right) \leq (1+\varepsilon)$$

$A(\varepsilon)$ runs in time polynomial in n (size of input) but not necessarily in $\frac{1}{\varepsilon}$

e.g $n^{5/\varepsilon}$

A fully polynomial approximation scheme is as above but $A(\varepsilon)$ is polynomial in both n and $\frac{1}{\varepsilon}$

e.g $\left(\frac{1}{\varepsilon}\right)^6 n^3$

Solving subset sum in exponential time:

$S = \{x_1, x_2, \dots, x_n\}$ with $x_1 \le x_2 \le \dots \le x_n$

$S_i = \{x_1, \dots, x_i\}$ $i = 0, 1, 2 \dots, n$

L_i = set of integers we can form as a sum of some elements from S_i

$L_0 = 0$, $L_1 = \{0, x_1\}$, $L_2 = \{0, x_1, x_2, x_1 + x_2\}$

remove numbers $> t$ from L_i

$L_i \to L_{i+1}$: add x_{i+1} to all numbers in L_i (removing values $> t$)

$L_0 \to L_1 \to L_2 \to \dots \to L_i \to L_{i+1} \to \dots \to L_n$

$\boxed{t^* \Leftarrow \text{max value in } L_n}$

list may double in each step so $|L_n|$ can be up to 2^n $\to$ algorithm is exponential

Example $S = \{1, 5, 7, 8\}$ $t = 19$

$L_0 = \{0\}$

$L_1 = \{0, 1\}$

$L_2 = \{0, 1, 5, 6\}$

$L_3 = \{0, 1, 5, 6, 7, 8, 12, 13\}$

$L_4 = \{0, 1, 5, 6, 7, 8, 9, 12, 13, 14, 15, 16\}$

If $S = \{1, 2, 4, 8, 16, \ldots 2^{n-1}\}$ and $t \geq 2^n$

$|L_i| = 2|L_{i-1}|$ for $i = 1, 2 \ldots n$ so running time

$O(2^n)$

Algorithm is polynomial if either

1. t is "small" as a function of n

2. all $x_i \leq p(n)$ for some polynomial p

Idea for approximation scheme:

δ-trimming the lists: represent some values by smaller ones

that is delete y if $\frac{y}{(1+\delta)} \le z \le y$

Example $\delta = \frac{1}{10}$ $\quad L = \{0, 10, 11, 12, 15, 17, 18, 20, 24, \dots\}$

trim $L \to L' = \{0, 10, 12, 15, 17, 20, 24, \dots\}$

Modified algorithm A'

Build lists as before, except that we trim L_i before building L_{i+1} (from L_i')

<u>Theorem</u> A' is a fully polynomial apx scheme for subset sum when we use trim factor $\delta = \frac{\varepsilon}{2n}$ $[0 < \varepsilon < 1$ is fixed$]$

<u>Proof</u>: We need to show that $\frac{y^*}{z^*} \le (1+\varepsilon)$ when

y^* is value of the optimum solution and

z^* is output from A' (largest number in L_n')

If we removed y from L_j in step j, then L_j' contains a value z s.t $\frac{y}{(1+\delta)} \le z \le y$

If z is later removed in step $j' > j$ then z is represented in $L_{j'}'$ by some z' with $\frac{z}{1+\delta} \le z'$

$$\frac{y}{1+\delta} \leq z \quad \wedge \quad \frac{z}{1+\delta} \leq z'$$

$\Downarrow$

$$z' \geq \frac{z}{1+\delta} \geq \frac{\frac{y}{1+\delta}}{1+\delta} = \frac{y}{(1+\delta)^2}$$

In particular at the last iteration y^* is represented by some element $z \in L_n'$ where

$$\frac{y^*}{(1+\delta)^n} \leq z \leq y^*$$

$z \in L_n'$ so $z \leq z^*$ implying $\frac{y^*}{(1+\delta)^n} \leq z \leq z^*$

so $\frac{y^*}{z^*} \leq (1+\delta)^n = \left(1+\frac{\varepsilon}{2n}\right)^n$

Calculus: $\left(1+\frac{\varepsilon/2}{n}\right)^n \leq e^{\varepsilon/2} = 1+\frac{\varepsilon}{2}+\left(\frac{\varepsilon}{2}\right)^2+\ldots \leq 1+\varepsilon$

when $0<\varepsilon<1$

Hence $\frac{y^*}{z^*} \leq (1+\varepsilon)$

$\Rightarrow$ A' is a $(1+\varepsilon)$-apx alg

Running time:

$L_n' = \{0, a_1, a_2, \dots a_{r+1}\}$

- $a_{i+1} > \left(1+\frac{\varepsilon}{2n}\right) a_i \quad \forall i$ as a_{i+1} was not removed when trimming
- Hence $\left(1+\frac{\varepsilon}{2n}\right)^r a_1 < a_{r+1} \leq t$

$$\Downarrow$$

$$r + \log_{\left(1+\frac{\varepsilon}{2n}\right)}(a_1) < \log_{\left(1+\frac{\varepsilon}{2n}\right)}(t)$$

$$\Downarrow$$

$$r < \log_{\left(1+\frac{\varepsilon}{2n}\right)}(t) = \frac{\ln t}{\ln\left(1+\frac{\varepsilon}{2n}\right)} \qquad \left(\log_a(x) = \frac{\ln x}{\ln a}\right)$$

So
$$|L_n'| = r+2 < 2 + \frac{\ln t}{\ln\left(1+\frac{\varepsilon}{2n}\right)}$$

$$\leq 2 + \frac{2n\left(1+\frac{\varepsilon}{2n}\right)\ln t}{\varepsilon} \qquad \left(\frac{x}{1+x} \leq \ln(1+x) \Rightarrow \frac{1+x}{x} \geq \frac{1}{\ln(1+x)}\right) \quad x = \frac{\varepsilon}{2n}$$

$$\leq 2 + \frac{3n \ln t}{\varepsilon} \qquad \left(1+\frac{\varepsilon}{2n} < \frac{3}{2}\right) \text{ as } 0<\varepsilon<1$$

$$= 3n \cdot \frac{1}{\varepsilon} \cdot \ln t + 2$$

polynomial in n, $\frac{1}{\varepsilon}$ and $\ln t$ $\square$