Section 22.7: Subset Sum Is NP-Hard

The Subset Sum Problem

Input: Positive integers $a_1, a_2, ..., a_n$, and a positive integer T.

Output: A subset of the a_i's that sum to T, $\left(S \subseteq \{1,2,...,n\} \text{ s.t. } \sum_{i \in S} a_i = T \right)$

Theorem: The independent set problem reduces to the subset sum problem.

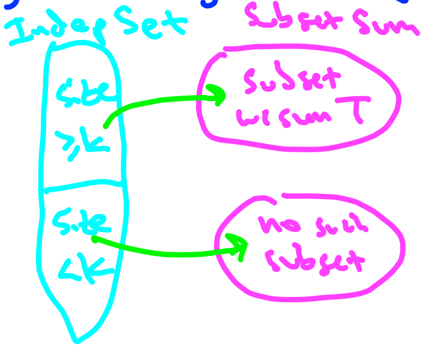
Corollary: The subset sum problem is NP-hard.

The Plan

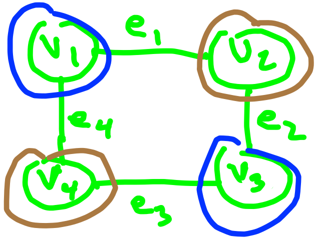
For now: just check for independent set of target size k.

Note: expect to use exponentially big numbers. [easy if numbers have polynomial magnitudes]

Key idea: use lower-order digits to encode incident edges.



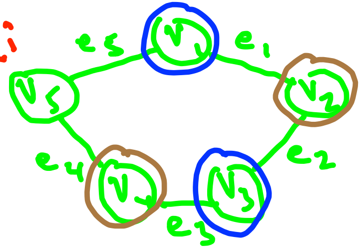
Ex#1:



$a_1 = 11{,}001$
$a_2 = 11{,}100$
$a_3 = 10{,}110$
$a_4 = 10{,}011$

sum = 21,111
sum = 21,111

Ex#2:



$a_1 = 110{,}001$
$a_2 = 111{,}000$
$a_3 = 101{,}100$
$a_4 = 100{,}110$
$a_5 = 100{,}011$

sum = 211,101
sum = 211,110

$b_1 = 10{,}000$
$b_2 = 1{,}000$
$b_3 = 100$
$b_4 = 10$
$b_5 = 1$

target T = 211,111

The Reduction

Given: instance $G=(V,E)$ of independent set. ($V=\{v_1,v_2,\ldots,v_n\}$, $E=\{e_1,e_2,\ldots,e_m\}$)

(A_i = edges incident to v_i)

Construct: $a_i = 10^m + \sum_{e_j \in A_i} 10^{m-j}$ for all v_i

$b_j = 10^{m-j}$ for all e_j.

For $k = n, n-1, n-2, \ldots, 2, 1$: [could also use binary search]

- Set $T = k \cdot 10^m + \sum_{j=1}^{m} 10^{m-j}$ [k in leading digits, 1's in m trailing digits]
- Compute a subset of a_i's & b_j's with sum T (if one exists) using assumed subroutine
- given such a subset, return the corresponding v_i's

The Reduction: Correctness Proof

(1) for any subset of a_i's and any number of b_j's, sum (written base-10) is digits of s followed by m digits in {0,1,2,3}

(2) has target sum T (k followed by m 1s) $\Leftrightarrow$ has k a_i's & the corresponding v_i's are an independent set of G.

Case 1: if G has no size-k independent set $\Rightarrow$ no subset with target sum. (+ subroutine will discover this)

Case 2: if G has a size-k independent set $\Rightarrow$ subroutine will return subset with sum T $\Rightarrow$ reduction will extract size-k independent set.

QED!