Set Cover

Abhiram Ranade

February 16, 2016

(Minimum cost) Set cover

Input: Collection C of sets $S_1, ..., S_m$
weight w_i for each set S_i.

Output: Subcollection $C' \subseteq C$ such that

- $\cup_{S \in C'} S = \sum_{S \in C} S$ $= U$
- $\sum_{S_i \in C'} w_i$ is as small as possible.

Weight of a set = "Cost"

Goal "buy" elements by "buying" sets

No polytime algorithm known. Many expect it does not exist.

Greedy algorithm

GreedySetcover(C, U){

1. Mark all elements of U "uncovered".
2. $C' =$ null.
3. While some elements remain uncovered {

   For each $S_j \notin C'$

   compute $u_j = w_j/$number of uncovered elements in S_j.

   Pick the set S_k that has minimum u_k.

   $C' = C' \cup \{S_k\}$.

   Mark the uncovered elements in S_k covered.

   }
4. Return C'

}

An example

$U = \{0, \ldots, 2^n - 1\}$.

$S_i = \{2^i, \ldots, 2^{i+1} - 1\}$, for $i = 0, \ldots, n-1$. $\quad w_i = 1$

$S_n = \{0, 2, 4, \ldots, 2^n - 2\}$ $\quad w_n = 1 + \epsilon$

$S_{n+1} = \{1, 3, 5, 7, \ldots, 2^n - 1\}$ $\quad w_{n+1} = 1 + \epsilon$

Analysis

C^* : Optimal collection OPT : optimal cost

U_i : set of elements remaining uncovered after iteration i of greedy algorithm $U_0 = U$

Renumber sets so that S_i = set picked by greedy in iteration i.

Lemma: $|U_i| \leq |U_{i-1}|(1 - w_i/OPT)$

Theorem: Greedy solution has weight $\leq OPT \cdot (1 + \ln |U|)$
Proof: Use $1 - x \leq \exp(-x)$, with equality only for $x = 0$.
$|U_i| \leq |U_{i-1}| \exp(-w_i/OPT)|U_{i-2}| \exp(-(w_i + w_{i-2})/OPT$
$\leq |U| \exp(w_1 + \ldots + w_i/OPT)$ < 1 at termination.
As soon as $w_1 + \ldots + w_i > OPT \cdot \ln |U|$.
Last iteration can add at most OPT. $\square$