CS 512: Set Covers

16:198:512
Instructor: Wes Cowan

Name:
RUID:

Consider a set S of n elements, and a collection of subsets $C = \{S_1, S_2, \ldots, S_m\}$ such that $S_i \subset S$ and

$$\bigcup_{i=1}^{m} S_i = S, \tag{1}$$

i.e., the collection of subsets cover all n elements of S. The problem of finding a *set cover* is to find a subset of C that covers every element.

The *Greedy Set Cover Algorithm* proceeds by selecting one of the S_i at a time, at each time selecting the S_i that covers *the most currently uncovered elements*. This very natural, simple algorithm, is in many ways as good as simple algorithms get for this problem. We have the following result that if k_{greedy} is the number of sets in the greedily constructed set cover and k^* is the size of a *minimal* set cover, we have

$$k_{\text{greedy}} \leq k^* \ln n. \tag{2}$$

This is a valuable result in that it tells us the greedy set cover is never *too* bad, compared to the minimum. But it is in many ways uninformative, especially since we have no particular way of knowing how large the minimum set cover is.

And unfortunately, this is a problem that largely can't be rectified. The issue is that the statement of the problem is so general, it is almost impossible to state anything specific or useful about the minimum set cover in general.

1) Argue that in any situation, $k^* \leq \min(n, m)$. *What does the problem look like if $m > n$? If $m \leq n$?)*
2) For any m, n, give an example of some $\{S_1, \ldots, S_m\}$ such that $k^* = 1$.
3) For any m, n, give an example of some $\{S_1, \ldots, S_m\}$ such that $k^* = m$, and one where $k^* = n$.

We could consider adding additional restrictions or structure to the problem in the following way: assume that every S_i is of the same fixed size, for instance $|S_i| = c$.

5) In this case, what is the smallest that m can be? What is the largest?
6) Argue that $k^* \geq \lceil n/c \rceil$. Give an example of some $\{S_1, \ldots, S_m\}$ that realizes this bound.
7) Argue that $k^* \leq n - (c - 1)$ *(consider small example c)*. Give an example of some $\{S_1, \ldots, S_m\}$ that realizes this bound.
8) For (7), argue that adding *any* subsets of size c to the collection not already present will strictly decrease the minimum set cover size.
9) Relate the answers to (6,7,8) to graphs and trees.

Note that in any of the examples you constructed above, if you take $m = k^*$ (i.e., there are no superfluous sets), then the greedy algorithm will return the same set cover as the minimal set cover - because there is only one set cover. So suppose that $k^* < m$.

10) In general, under the restriction that $|S_i| = c$, how much worse than the minimal set cover can the greedy set cover be?