Due: Tuesday, May 9, 2023, 11:59pm

Submit online at (clickable link): `https://forms.gle/1eAmuDcgASBFocbG6`

1 Set Cover Example

Consider the following Set Cover example from the textbook. The underlying set X is a set of 12 characters:

$$X = \{\texttt{a}, \texttt{d}, \texttt{e}, \texttt{h}, \texttt{i}, \texttt{l}, \texttt{n}, \texttt{o}, \texttt{r}, \texttt{s}, \texttt{t}, \texttt{u}\}$$

The family $\mathcal{F}$ of subsets of X are depicted as words, but they are really just subsets of characters. The words are:

$$\mathcal{F} = \{\ \texttt{arid}, \texttt{dash}, \texttt{drain}, \texttt{heard}, \texttt{lost}, \texttt{nose}, \texttt{shun}, \texttt{slate}, \texttt{snare}, \texttt{thread}\ \}$$

For example, `shun` is really the subset $\{\texttt{h}, \texttt{n}, \texttt{s}, \texttt{u}\}$.

1. Which words would be chosen by the Greedy Set Cover algorithm for this instance of Set Cover? Break ties by choosing the word that is earlier in dictionary order.
2. Is the solution provided by Greedy Set Cover an optimum solution?

2 Bin Packing: First-Fit Decreasing

Instructions: Below is a proof that the First-Fit Decreasing (FFD) heuristic uses no more than $3/2 \cdot c^*$ bins, where c^* is the number bins used in the optimum packing. At various points in the proof, you will be asked to provide an explanation or analysis. The correct answers to these questions are short and direct. If you are writing a lot, you most likely missed the point of the question.

Recall that in the Bin Packing problem we have n items with sizes $s_1, s_2, \ldots, s_n$. We are asked to place the items into bins so that the sum of the sizes of the items in each bin is ≤ 1. The objective is to do so with as few bins as possible.

In the First-Fit Decreasing (FFD) heuristic, we sort the items by size. So, we can assume that

$$1 \geq s_1 \geq s_2 \geq s_3 \geq \cdots \geq s_n.$$

We take the items in order of decreasing size and place the current item in the first bin with enough remaining space to hold the current item. If none of the bins have enough space, then we start a new bin. The bins are considered in the order they were started — oldest bin first.

Let c^* be the number of bins used in the optimum bin packing. Let i_0 be the index of the first item that FFD places in bin number $c^* + 1$.

Task 1. There is no guarantee that i_0 exists. What happens if there is no such i_0?

Claim 1: $s_{i_0} \leq 1/2$.

Proof by contradiction: suppose $s_{i_0} > 1/2$. Then $s_1, s_2, \ldots, s_{i_0}$ are all $> 1/2$. So, there are $c^* + 1$ items that are $> 1/2$ which means the optimum bin packing must use at least $c^* + 1$ bins, a contradiction. QED.

Task 2. How do we know there are $c^* + 1$ items $> 1/2$?

Task 3. Why would the optimum bin packing use at least $c^* + 1$ bins?

Task 4. Why would the optimum bin packing using $c^* + 1$ bins be a contradiction?

Let c be the total number of bins used by FFD. We call the bins numbered $c^* + 1$ through c the "extra bins." Let t be the number of items that FFD placed in the extra bins and let $z_1, z_2, \ldots, z_t$ be their sizes.

Task 5. There are $n - i_0$ items that come after item i_0, the first item to be placed in an extra bin. Why isn't t just equal to $n - i_0 + 1$?

Claim 2: $t \leq c^* - 1$.

Proof by Contradiction: Suppose $t \geq c^*$. For $1 \leq j \leq c^*$, let b_j be the sum of the sizes of the items that FFD placed in bin number j. Then, for $1 \leq j \leq c^*$, $b_j + z_j > 1$. Thus,

$$\sum_{j=1}^{c^*} (b_j + z_j) > \sum_{j=1}^{c^*} 1 = c^*$$

Since the sum of the sizes of all of the items cannot exceed c^*, this is a contradiction. QED.

Task 6. Why is $b_j + z_j > 1$?

Task 7. Why can't the sum of all the sizes exceed c^*?

Finally, since each item that FFD placed in an extra bin has size $\leq 1/2$ and since there are no more than $c^* - 1$ of them, FFD uses no more than $\lceil (c^* - 1)/2 \rceil \leq c^*/2$ extra bins. Thus, FFD uses no more than $3/2 \cdot c^*$ bins in total.

Task 8. How do we know that FFD uses no more than $\lceil (c^* - 1)/2 \rceil$ extra bins?

Task 9. Why is the total number of bins used by FFD less than or equal to $3/2 \cdot c^*$?