

Greedy Heuristics for set cover

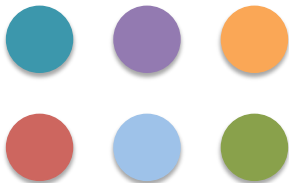
Eirini Asteri, Jessica Hoffmann

University of Texas, Austin

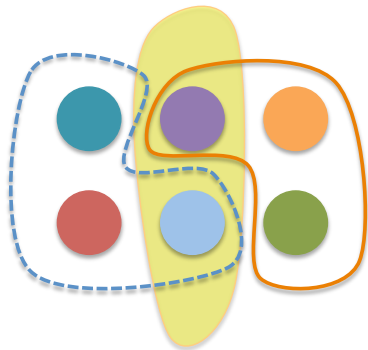
Outline

Set Cover Problem & Greedy Algorithm

- Number of distinct elements e_i

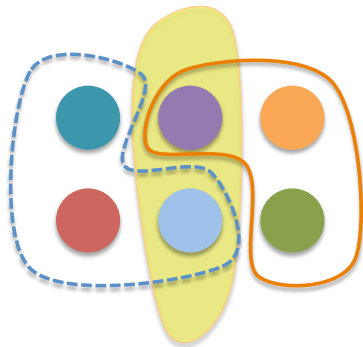


Set Cover Problem & Greedy Algorithm



- Number of distinct elements e_i
- Number of sets S_j

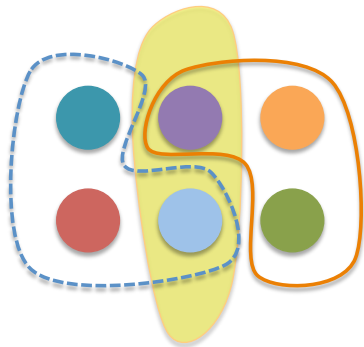
Set Cover Problem & Greedy Algorithm



- Number of distinct elements e_i
- Number of sets S_j
- A weight for each set w_j



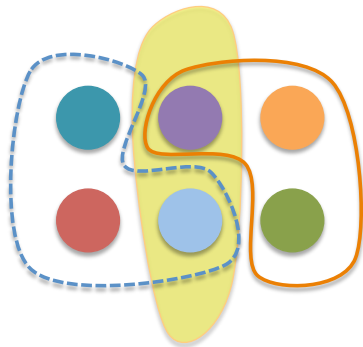
Set Cover Problem & Greedy Algorithm



- Number of distinct elements e_i
- Number of sets S_j
- A weight for each set w_j
- **Goal** Find set cover with minimum weight



Set Cover Problem & Greedy Algorithm



- Number of distinct elements e_i
- Number of sets S_j
- A weight for each set w_j
- **Goal** Find set cover with minimum weight
- **Greedy Choice**

$$best_set = \arg \min_i \frac{w_i}{\hat{S}_i}$$



Basic Preprocessing

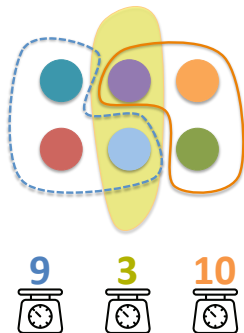
Get rid of redundant sets 🗑️

If $S_{small} \subseteq S_{big}$ and $\text{weight}(S_{small}) \geq \text{weight}(S_{big})$ then S_{small} is a redundant set.

Theoretical Analysis

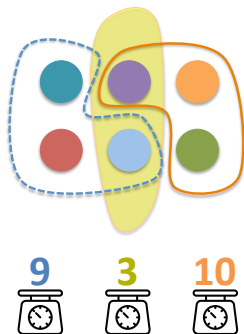
Heuristics [Intuition]

- Elements with frequency 1 should be covered first

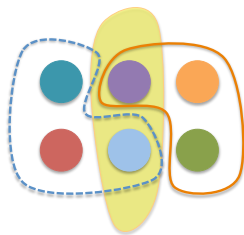


Heuristics [Intuition]

- Elements with frequency 1 should be covered first
- Extend idea to **“infrequent”** elements



Heuristics [Intuition]

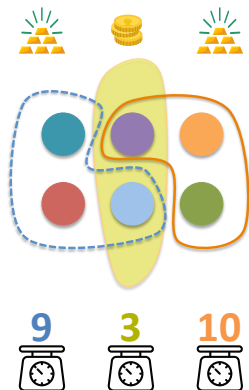


- Elements with frequency 1 should be covered first
- Extend idea to **“infrequent”** elements
- Assign a value to each element

$$\text{value}(\text{element}) = \frac{1}{\text{frequency}-1}$$



Heuristics [Intuition]



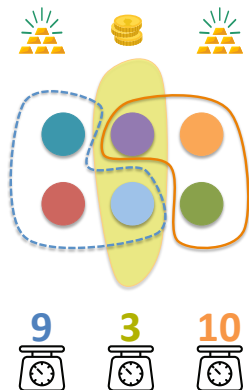
- Elements with frequency 1 should be covered first
- Extend idea to **“infrequent”** elements
- Assign a value to each element

$$\text{value}(\text{element}) = \frac{1}{\text{frequency}-1}$$

- Assign value to each set

$$\text{value}(\text{set}) = \sum \text{value}(\text{element})$$

Heuristics [Intuition]



- Elements with frequency 1 should be covered first
- Extend idea to “**infrequent**” elements
- Assign a value to each element

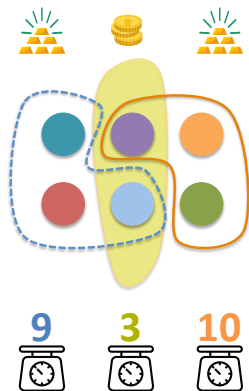
$$\text{value}(\text{element}) = \frac{1}{\text{frequency}-1}$$

- Assign value to each set

$$\text{value}(\text{set}) = \sum \text{value}(\text{element})$$

- Choose a set with **small weight** and **large value**!

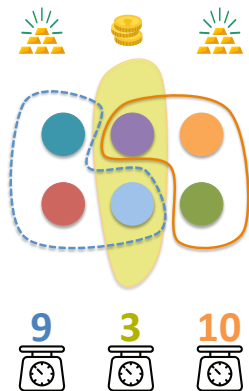
Heuristics, General Framework



- New Greedy Choice

$$best_set = \arg \min_j \frac{w_j}{v_j} = \arg \min_j \frac{w_j}{\sum \text{value}(e_i)}$$

Heuristics, General Framework




- New Greedy Choice

$$best_set = \arg \min_j \frac{w_j}{v_j} = \arg \min_j \frac{w_j}{\sum \text{value}(e_i)}$$

- Regular Greedy is a Special Case

$$\text{if } \text{value}(e_i) = 1 \rightarrow \frac{w_j}{\sum \text{value}(e_i)} = \frac{w_j}{|\hat{S}_j|}$$

Another Heuristic Valuation..

Dig Deeper, Extract more Information 

$$\text{value}(e_i) = \frac{\sum_{S_j: e_i \in S_j} \text{average_weight}(S_j)}{\text{frequency}(e_i) - 1}$$

An element is valuable if it is contained in “expensive” sets.

[**Intuition**] Choose a “**cheap**” set that contains elements that are “**expensive in the market**”.

Value Functions Tested

- Greedy: $v(e_i) = 1$
- H 1: $v(e_i) = \frac{1}{f_i - 1}$
- H 2: $v(e_i) = 1 + \frac{1}{f_i - 1}$
- H 3: $v(e_i) = \exp(-f_i)$
- H 4: $v(e_i) = \frac{|\hat{S}_j|}{f_i - 1}$
- H 7: $v(e_i) = \frac{1}{(f_i - 1)^2}$
- H 8: $v(e_i) = \frac{1}{(f_i - 1)^3}$
- H 9: $v(e_i) = \frac{1}{\sqrt{f_i - 1}}$
- H 10: $v(e_i) = \frac{\sum w_j / |\hat{S}_j|}{f_i - 1}$
- H 11:
$$v(e_i) = c + \frac{\sum w_j / |\hat{S}_j|}{f_i - 1}$$

Data Sets

Experimental Results

References