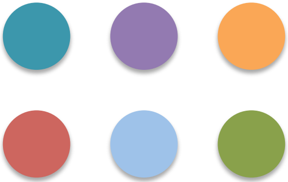
Greedy Heuristics for Set Cover

Eirini Asteri, Jessica Hoffmann

University of Texas, Austin

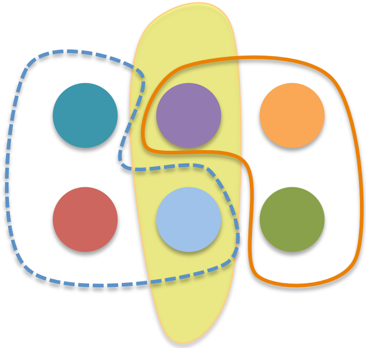
Outline

Set Cover Problem & Greedy Algorithm



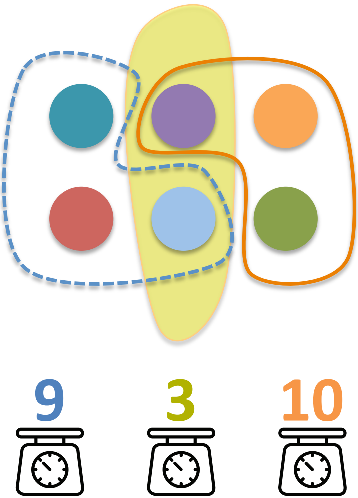
- Number of distinct elements e_i

Set Cover Problem & Greedy Algorithm



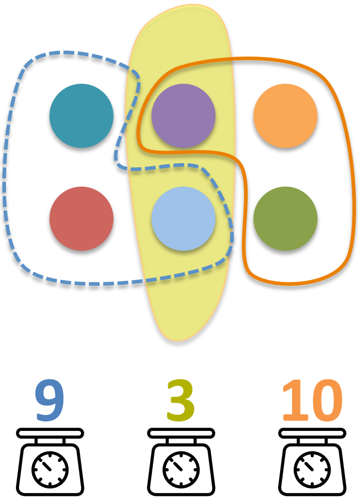
- Number of distinct elements e_i
- Number of sets S_j

Set Cover Problem & Greedy Algorithm



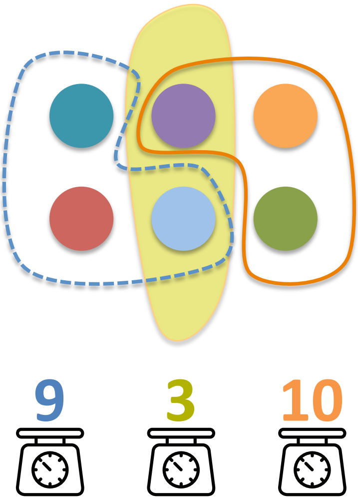
- Number of distinct elements e_i
- Number of sets S_j
- A weight for each set w_j

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

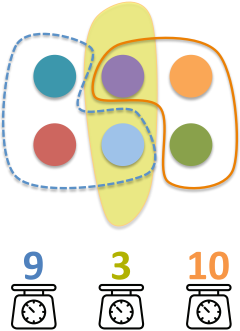
Theoretical Analysis

Approximation result

Any greedy heuristic, caracterized by its value() function, is a $\dfrac{\max\limits_{e_i} value(e_i)}{\min\limits_{e_i} value(e_i)} \cdot H_n$-approximation.

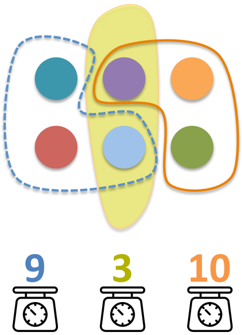
Remark: This shows any greedy heuristics will have worse theoretical guarantees if our analysis is tight.

Heuristics [Intuition]



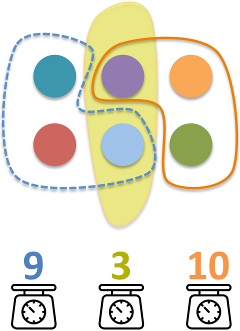
- Elements with frequency 1 should be covered first

Heuristics [Intuition]



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

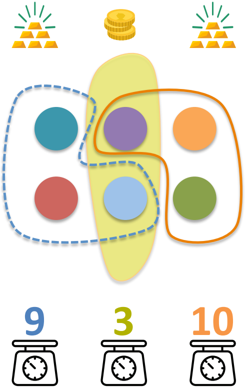
Heuristics [Intuition]



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

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

Heuristics [Intuition]



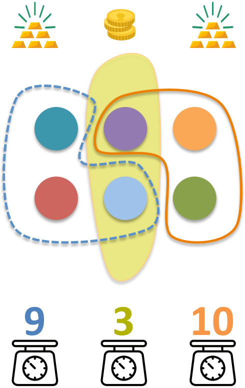
- Elements with frequency 1 should be covered first
- Extend idea to **“infrequent”** elements
- Assign a value to each element

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

- Assign value to each set

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

Heuristics [Intuition]



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

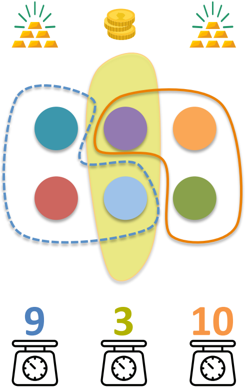
$$\text{value(element)} = \frac{1}{\text{frequency-1}}$$

- Assign value to each set

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

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

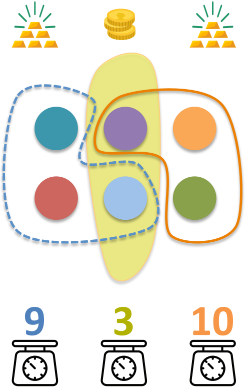
Heuristics, General Framework



- 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

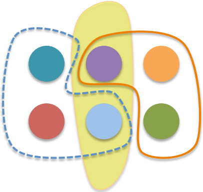
$$best\_set = \arg\min_j \frac{w_j}{v_j} = \arg\min_j \frac{w_j}{\sum \text{value}(e_i)}$$

- Regular Greedy is a Special Case

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

Another Heuristic Value Function..

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

  $v(e_i) = c + \dfrac{\sum w_j / |\hat{S}_j|}{f_i - 1}$

Data Sets

Experimental Results

References