

Problème de la couverture d'ensemble

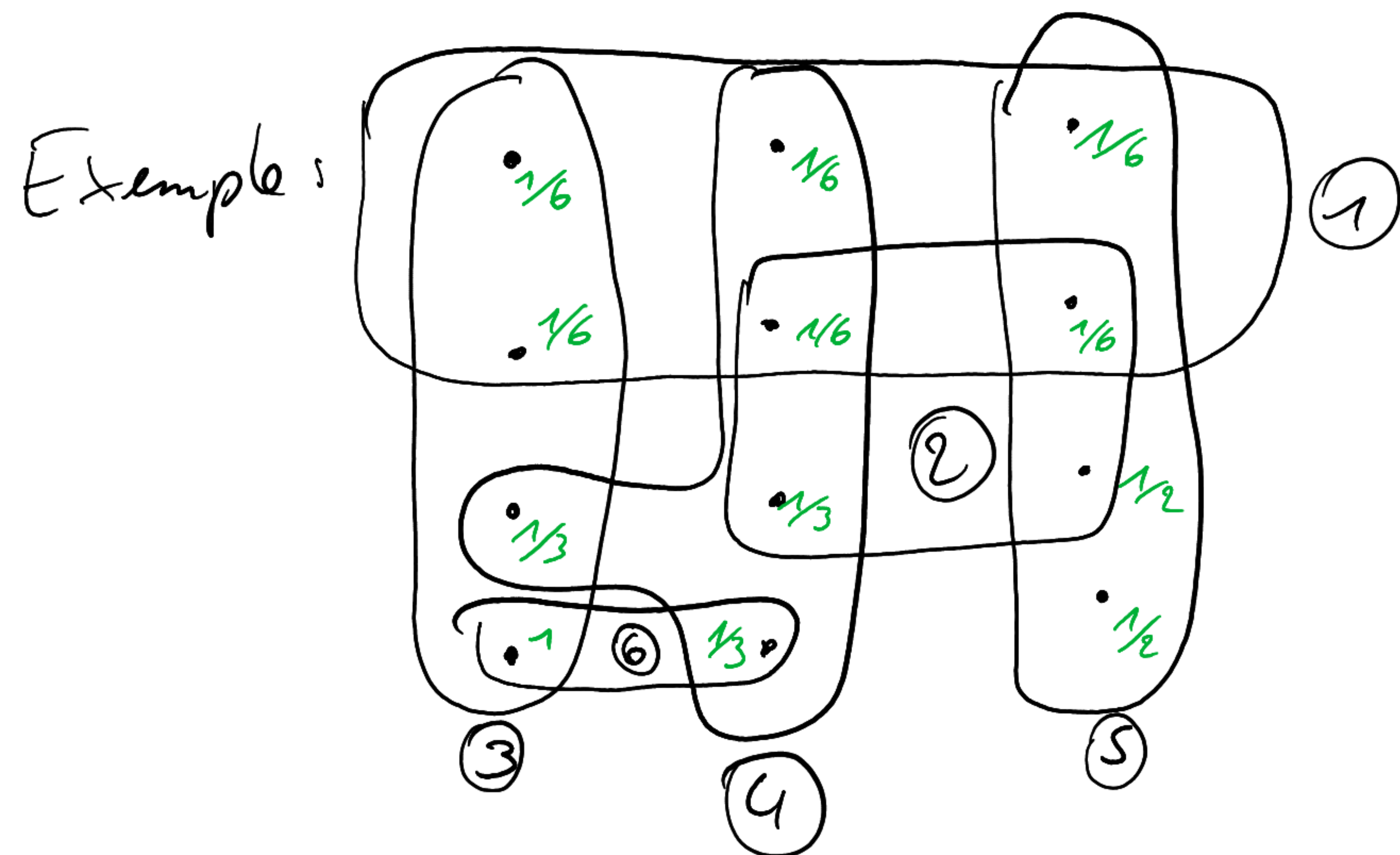
Définition: Soient X un ensemble et $\mathcal{C} \subseteq \mathcal{P}(X)$.

On dit que $\mathcal{C}$ **couvre** X si $\forall x \in X, \exists S \in \mathcal{C}, x \in S$.
($\Leftrightarrow \bigcup_{S \in \mathcal{C}} S = X$).

Problème: Couverture d'ensemble / SET COVER

Entrée: X ensemble fini et $\mathcal{F} \subseteq \mathcal{P}(X)$ tel que:
 $\mathcal{F}$ couvre X .

Sortie: $\mathcal{C} \subseteq \mathcal{F}$ minimal tel que $\mathcal{C}$ couvre X .



Exécution algo:
 $1 \rightarrow 4 \rightarrow 5 \rightarrow 6$

Couverture optimale: 3, 4, 5.

Propriété: SET COVER est NP-complet.

Fonction couverture $(X, \mathcal{F})$:

$$U = X$$

$$C = \emptyset$$

Tant que $U \neq \emptyset$:

Choisir un $S \in \mathcal{F}$ maximal pour $|S \cap U|$

$$U = U \setminus S$$

$$C = C \cup \{S\}$$

Retourner

Idee : On construit une couverture de sommets petit à petit en sélectionnant à chaque l'ensemble contenant le plus d'éléments non couverts.

Terminaison : Ok car $\mathcal{F}$ borne X

Complexité : $O(\underbrace{|X| \times |\mathcal{F}|}_{\text{boucle}} \times \underbrace{\min(|X|, |\mathcal{F}|)}_{\text{nb de tours}})$

Correction : On construit ici une $(\ln(|X|) + 1)$ -approximation.

Il faut montrer que $|C| \leq \underbrace{|C^*|}_{\text{optimal}} \times \ln(|X| + 1)$.

S_i : i^{e} sous-ensemble choisi par l'algorithme.

À chaque étape i , l'algorithme distribue 1 pour tout S_i réparti de façon équilibrée entre tous les éléments.

On note c_x le coût attribué à x , pour tout $x \in X$.

(Compléter l'exemple)

On a :
$$c_x = \frac{1}{|S_i \setminus (S_1 \cup S_2 \cup \dots \cup S_{i-1})|} \quad \text{si } c_x$$

a été découvert à l'étape i .

On a :
$$|C| = \sum_{x \in X} c_x.$$

On pose :
$$V^* = \sum_{S \in \mathcal{F}} \sum_{x \in S} c_x.$$

Comme chaque élément de X apparaît au moins une fois dans C^* , on a :
$$V^* \geq \sum_{x \in X} c_x.$$

Donc :
$$|C| \leq V^*.$$

Montrons que pour $S \in \mathcal{F}$,
$$\sum_{x \in S} c_x \leq H(|S|)$$
 avec $H(n) = \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n}$. ($H(0) = 0$).

On pose : $U_i = |S \setminus (S_1 \cup S_2 \cup \dots \cup S_{i-1})|$ avec $U_0 = |S|$, le nombre d'éléments de S non couverts après i étapes.

On a : $|S_i \setminus (S_1 \cup \dots \cup S_{i-1})| \geq |S \setminus (S_1 \cup \dots \cup S_{i-1})| = U_{i-1}$ car sinon l'algorithme aurait choisi S et non S_i .

Avec $k = \arg \min (U_k = 0)$, on a :

$$\sum_{x \in S} c_x = \sum_{i=0}^k (U_{i-1} - U_i) \times \frac{1}{|S_i \setminus (S_1 \cup \dots \cup S_{i-1})|} \quad (\text{nb éléments couverts} \times \text{coût}).$$

$$\text{Dac: } \sum_{x \in S} c_x \leq \sum_{i=1}^k (u_{i-1} - u_i) \times \frac{1}{u_{i-1}}$$

Ne pas
hésiter à passer

$$\left\{ \begin{aligned} &= \sum_{i=1}^k \sum_{j=u_i+1}^{u_{i-1}} \frac{1}{u_{i-1}} \\ &\leq \sum_{i=1}^k \sum_{j=u_i+1}^{u_{i-1}} \frac{1}{j} \quad \text{car } j \leq u_{i-1} \\ &= \sum_{i=1}^k \left(\sum_{j=1}^{u_{i-1}} \frac{1}{j} - \sum_{j=1}^{u_i} \frac{1}{j} \right) \\ &= \sum_{i=1}^k (H(u_{i-1}) - H(u_i)) \\ &= H(u_0) - \underbrace{H(u_k)}_{= H(0) = 0} \\ &= H(u_0) \\ &= H(|S|) \end{aligned} \right.$$

$$\text{Dac: } \sum_{x \in S} c_x \leq H(|S|), \quad \forall S \in \mathcal{F}.$$

$$\text{Dac: } \sum_{S \in C^*} \sum_{x \in S} c_x = v^* \leq \sum_{S \in C^*} H(|S|)$$

$$v^* \leq |C^*| \times \underbrace{H(\max(|S|, S \in \mathcal{F}))}_{\leq |X|}$$

$$\text{Dac: } |C| \leq |C^*| \times H(|X|) \leq |C^*| \times (\ln(|X|) + 1)$$