In the <u>set cover problem</u>, we are given a ground set of elements $E = \{e_1, \dots, e_n\}$, some subsets of those elements $S_1, S_2, \dots, S_m$ where each $S_j \subseteq E$, and a nonnegative weight $w_j \geq 0$ for each subset S_j. The goal is to find a minimum-weight collection of subsets that covers all of E; that is, we wish to find an $I \subseteq \{1, \dots, m\}$ that minimizes $\sum_{j \in I} w_j$ subject to $\bigcup_{j \in I} S_j = E$. If $w_j = 1$ for each subset j, the problem is called the unweighted set cover problem.

The vertex cover problem is a special case of the set cover problem. For any instance of the vertex cover problem, create an instance of the set cover problem in which the ground set is the set of edges, and a subset S_i of weight w_i is created for each vertex $i \in V$ containing the edges incident to i. For any vertex cover C, there is a set cover $I = C$ of the same weight, and vice versa.

<u>ILP formulation</u>:

$$\text{minimize} \sum_{j=1}^{m} w_j x_j$$

$$\text{subject to} \sum_{j : e_i \in S_j} x_j \geq 1, \quad i = 1, \dots, n$$

$$x_j \in \{0, 1\}, \quad j = 1, \dots, m$$

<u>LP relaxation</u>:

$$\text{minimize} \sum_{j=1}^{m} w_j x_j$$

$$\text{subject to} \sum_{j : e_i \in S_j} x_j \geq 1, \quad i = 1, \dots, n,$$

$$x_j \geq 0, \quad j = 1, \dots, m$$

Let Z^*_{IP} denote the optimum value of the ILP.
We have $Z^*_{IP} = OPT$, where OPT is the value of an optimum solution to the set cover problem.
Let Z^*_{LP} denote the optimum value of the LP.
We have: $Z^*_{LP} \leq Z^*_{IP} = OPT$.
Let x^* denote an optimal solution to the LP. We include subset S_j in our solution if and only if $x_j^* \geq 1/f$, where f is the maximum number of sets in which any element appears. Let $f_i = |\{j : e_i \in S_j\}|$ be the number of sets in which element e_i appears, $i = 1, \dots, n$; then $f = \max_{i=1,\dots n} f_i$. Let I denote the indices j of the subsets in this solution.

The collection of subsets S_j, $j \in I$, is a set cover:
We will show that each element e_i is covered. Because the optimal solution x^* is a feasible solution to the linear program: $\sum_{j : e_i \in S_j} x_j^* \geq 1$ for element e_i. By the definition of f_i and of f, there are $f_i \leq f$ terms in the sum, so at least one term must be at least $1/f$. Thus, for some j such that $e_i \in S_j$, $x_j^* \geq 1/f \Rightarrow$ $j \in I$, and element e_i is covered.

The LP-rounding algorithm is an f-approximation algorithm for the set cover problem:
The LP can be solved in polynomial time using ellipsoid algorithm. By our construction of I: $1 \leq f \cdot x_j^*$, $\forall j \in I$.

$$\sum_{j \in I} w_j \leq \sum_{j=1}^{m} w_j (f \cdot x_j^*) = f \sum_{j=1}^{m} w_j x_j^* = f \cdot Z^*_{LP} \leq f \cdot OPT$$

$\Rightarrow \underline{\left(\sum_{j \in I} w_j \right) / OPT \leq f}$. In the special case of the vertex cover problem, $f_i = 2$ $\forall i \in V$ giving a 2-approx. algo.

Primal LP:

minimize $\sum_{j=1}^{n} c_j x_j$

subject to $\sum_{j=1}^{n} a_{ij} x_j \geq b_i, \quad i = 1, \ldots, m$

$x_j \geq 0, \quad j = 1, \ldots, n$

Dual LP: We try to find a lower bound for the optimal primal solution. For this we multiply the m constraints by values (non-negative) $y_i (1 \leq i \leq m)$. We add the constraints. In order to get a lower bound of the optimal primal solution, the coefficient of x_j in the sum should be bounded above by c_j. We are trying to find the best possible lower bound in this way. This gives rise to the dual LP as follows:

maximize $\sum_{i=1}^{m} b_i y_i$

subject to $\sum_{i=1}^{m} a_{ij} y_i \leq c_j, \quad j = 1, \ldots, n$

$y_i \geq 0, \quad i = 1, \ldots, m$

LP-duality theorem: The primal program has finite optimum if and only if it's dual has finite optimum. Moreover, if $x^* = (x_1^*, \ldots, x_n^*)$ and $y^* = (y_1^*, \ldots, y_m^*)$ are optimal solutions for the primal and dual programs, respectively, then

$$\sum_{j=1}^{n} c_j x_j^* = \sum_{i=1}^{m} b_i y_i^*.$$

<u>Weak duality theorem</u>: If $x=(x_1,\dots,x_n)$ and $y=(y_1,\dots,y_m)$ are feasible solutions for the primal and dual program, respectively, then

$$\sum_{j=1}^{n} c_j x_j \geq \sum_{i=1}^{m} b_i y_i .$$

$$\sum_{j=1}^{n} c_j x_j \geq \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij} y_i\right) x_j \quad (y \text{ is dual feasible}, x_j \geq 0)$$

$$= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij} x_j\right) y_i \geq \sum_{i=1}^{m} b_i y_i \quad (x \text{ is primal feasible}, y_i \geq 0).$$

<u>Complementary slackness conditions</u>: Let x and y be primal and dual feasible solutions, respectively. Then, x and y are both optimal if and only if all of the following conditions are satisfied:

<u>Primal complementary slackness conditions</u>:

For each $1 \leq j \leq n$: either $x_j = 0$ or $\sum_{i=1}^{m} a_{ij} y_i = c_j$; and

<u>Dual complementary slackness conditions</u>:

For each $1 \leq i \leq m$: either $y_i = 0$ or $\sum_{j=1}^{n} a_{ij} x_j = b_i$

<u>Dual LP for Set Cover</u>:

$$\text{maximize} \sum_{i=1}^{n} y_i$$

$$\text{subject to} \sum_{i: e_i \in S_j} y_i \leq w_j, \quad j = 1, \dots, m,$$

$$y_i \geq 0, \quad i = 1, \dots, n$$