Albert-Ludwigs-Universität Freiburg
Institut für Informatik
Prof. Dr. Susanne Albers
Dr. Alexander Souza

Randomisierte Algorithmen und probabilistische Methoden

Aufgabe 1 Set Cover

Beim Problem Set Cover ist eine n-elementige Menge U und eine Menge S von Teilmengen über U gegeben, d.h. $S = \{S_1, \dots S_m\}$ mit $S_i \subseteq U$. Jeder Teilmenge S_i sind Kosten $c(S_i) \geq 0$ zugeordnet. Die Aufgabe besteht darin, eine Teilmenge $C \subseteq S$ zu finden, so dass $\cup_{S_i \in C} S_i = U$ und die Gesamtkosten $c(C) = \sum_{S_i \in C} c(S_i)$ minimal sind. Ohne Einschränkung setzen wir daher voraus, dass $\cup_{S_i \in S} S_i = U$ gilt, da das Problem ansonsten nicht lösbar ist. Set Cover ist NP-vollständig.
Das Ziel dieser Übung ist es einen Algorithmus dafür zu finden, der die Technik des Randomisierten Rundens verwendet und dessen Lösung im Erwartungswert höchstens $O(\log n)$ mal so teuer ist, wie die optimale Lösung. Gehen Sie dabei wie folgt vor:

(1) Vergewissern Sie sich, dass folgendes lineare ganzzahlige Programm (IP) äquivalent zu Set Cover ist:

$$\sum_{i=1}^{m} c(S_i)x_i \longrightarrow \min$$
$$\sum_{i:j \in S_i} x_i \geq 1 \text{ für alle } j \in U$$
$$x_i \in \{0, 1\}$$

(2) Der Algorithmus wiederholt das Randomisierte Runden aus der Vorlesung t mal und gibt die Vereinigung C der t ermittelten Lösungen aus.

(3) Zeigen Sie nun folgendes:

(a) Zeigen Sie, dass die Wahrscheinlichkeit, dass ein festes $j \in U$ bei einer der t Wiederholungen *nicht* überdeckt wird höchstens $\frac{1}{e}$ ist.

(b) Wie wahrscheinlich ist es, dass dieses j in allen t Wiederholungen nicht überdeckt wird?

(c) Zeigen Sie, dass $t = \lceil \log(4n) \rceil$ Wiederholungen genügen, um die Wahrscheinlichkeit, dass der Algorithmus keine gültige Lösung bestimmt mit $\frac{1}{4}$ beschränken zu können.

(d) Warum gilt $\mathbb{E}[c(C)] = O(\log n)\,\mathrm{OPT}_{LP}$, wobei OPT_{LP} den Wert der optimalen Lösung des linearen Programms bezeichnet.

(e) Argumentieren Sie abschliessend, warum OPT_{LP} eine untere Schranke für den optimalen Wert OPT_{IP} obigen ganzzahligen linearen Programms ist.