Prof. Dr. M. Grohe E. Fluck, A. Riazsadri, J. Feith

Übungsblatt 11 mit Lösungen

Abgabetermin: Mittwoch, der 25. Januar 2023 um 14:30

Aufgabe 4 (Set Packing) **6 Punkte**

Wir betrachten das folgende Entscheidungsproblem:

SET PACKING

Eingabe: Eine Menge U, eine Menge $\mathcal{S} \subseteq \mathrm{Pot}(U)$, eine Zahl $k \in \mathbb{N}$.

Frage: Enthält $\mathcal{S}$ k disjunkte Mengen, d.h. existiert eine Menge $\mathcal{S}' = \{S_1, \dots, S_k\} \subseteq \mathcal{S}$ der Größe k mit $S_i \cap S_j = \emptyset$ für alle $i \neq j$?

Zeigen Sie, dass SET PACKING NP-vollständig ist.

Lösung:

Für eine gegebene SET PACKING-Instanz $(U, \mathcal{S} = \{S_1, \dots, S_m\}, k)$. Um eine Auswahl von k Mengen in $\mathcal{S}$ zu kodieren, nutzen wir das Zertifikat

$$C(i_1, \dots, i_k) = \mathrm{bin}(i_1)\# \dots \# \mathrm{bin}(i_k),$$

wobei $1 \leq i_1, \dots, i_k \leq m$ die Indizes der ausgewählten Mengen sind. Offenbar ist die Länge des Zertifikates durch $O(k \log m) = O(n^2)$ beschränkt, wobei n die Kodierungslänge der Eingabe ist.

Um ein gegebenes Zertifikat zu verifizieren, gehen wir wie folgt vor:

1. Teste, ob $C = C(i_1, \dots, i_\ell)$ für paarweise verschiedene Indizes $1 \leq i_1, \dots i_\ell \leq m$ mit $\ell \geq k$ ist.

2. Prüfe, ob die Mengen in $\mathcal{S}$ mit den Indizes aus C paarweise verschieden sind und verwerfe, falls dies nicht der Fall ist.

3. Sonst akzeptiere.

Die Laufzeit des Verifizierers ist offensichtlich polynomiell in der Eingabegröße. Weiterhin akzeptiert der Verifizierer genau dann, wenn das Zertifikat mittels der oben beschriebenen Syntax eine korrekte Lösung der SET PACKING-Instanz $(U, \mathcal{S}, k)$ kodiert. Somit folgt SET PACKING $\in$ NP.

Als nächstes zeigen wir INDEPENDENT SET $\leq_p$ SET PACKING, woraus auch die NP-Schwere von SET PACKING folgt.

Sei dazu (G, k) eine INDEPENDENT SET-Instanz. Wir definieren $E_v(G) = \{e \in E(G) \mid v \text{ ist inzident zu } e\}$ und

$$\mathcal{S}(G) = \{E_v(G) \mid v \in V(G)\}.$$

Dann sind für zwei Knoten v und w die Mengen E_v und E_w verschieden, solange v und w nicht isolierte Knoten sind oder vw keine isolierte Kante ist (d.h. die Kante vw formt eine Zusammenhangskomponente von G) Daher "bereinigen" wir G zunächst um die Menge I_V

aller isolierten Knoten und die Menge I_E aller isolierten Kanten (samt ihrer Endpunkte). Sei G' der so erhaltene Graph ohne isolierte Knoten, und sei $k' = k - |I_V| - |I_E|$. Dann hat G' ein Independent Set der Größe k' genau dann, wenn G ein Independent Set der Größe k hat, da wir in G zu jedem IS von G' alle Knoten in I_V und je einen Endpunkt einer jeden Kante in I_E hinzufügen können.

Nun sind alle Mengen in $\mathcal{S}(G')$ verschieden und es gilt $E_v(G') \cap E_w(G') = \emptyset$ genau dann, wenn $vw \notin E(G')$.

Sei X ein Independent Set der Größe k' in G'. Dann sind auch die Mengen $E_X = \{E_x(G') \mid x \in X\} \subseteq \mathcal{S}$ paarweise disjunkt, und es $|E_X| = k'$.

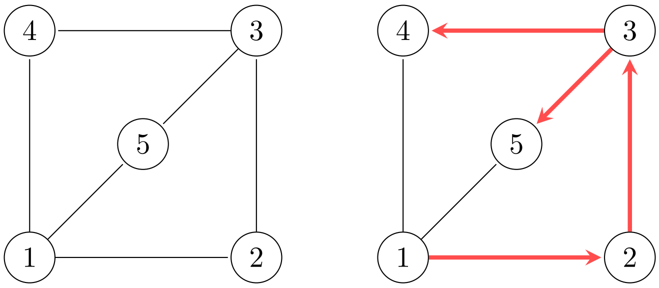
Sei nun andersherum $E_X \subseteq \mathcal{S}$ eine Menge von k' paarweise disjunkter Mengen. Dann gilt $E_X = \{E_x(G') \mid x \in X\}$ für eine Menge $X \subseteq V(G')$ der Größe k', welche folglich ein Independent Set von G' der Größe k' sind.

Da (G', k') und folglich $(E(G'), \mathcal{S}(G'), k')$ in Polynomialzeit aus (G, k) berechnet werden kann, folgt Independent Set $\leq_p$ Set Packing. Somit ist Set Packing NP-schwer.

Aufgabe 5 (Approximation für Vertex Cover) **9 (2 + 3 + 2 + 2) Punkte**

Wir erinnern uns an die Tiefensuche (DFS) in Graphen. Ein Durchlauf der Tiefensuche auf einem zusammenhängenden Graphen G kann durch einen (gerichteten) *DFS-Spannbaum* T repräsentiert werden, welcher genau die Kanten enthält, die in der Ausführung "genommen" wurden.

Zum Beispiel repräsentiert der rot (und dick) gedruckte Spannbaum T rechts einen DFS-Durchlauf auf dem linken Graphen, beginnend beim Knoten 1.



Wir bezeichnen die Knoten eines solchen Baums T, welche keine ausgehende Kante besitzen, als die *Blätter* von T; alle anderen Knoten nennen wir *interne Knoten*.

a) Sei G ein zusammenhängender Graph und T ein DFS-Spannbaum von G. Zeigen Sie, dass die Blätter von T ein Independent Set von G sind.

b) Zeigen Sie, dass jeder Baum T ein Matching besitzt, welches alle internen Knoten von T überdeckt.

c) Folgern Sie, dass das folgende Verfahren ein 2-Approximationsalgorithmus für Vertex Cover auf zusammenhängenden Graphen ist:

1. Wähle einen beliebigen Startknoten $r \in V(G)$.
2. Berechne mittels Tiefensuche von r aus einen DFS-Spannbaum T von G.
3. Gebe $I(T) = \{v \in V(G) \mid v \text{ ist interner Knoten von } T\}$ aus.

d) Zeigen Sie durch die Angabe einer geeigneten Familie $\mathcal{G} = (G_n)_{n\in\mathbb{N}}$ von Graphen, dass für alle $\varepsilon > 0$ das obrige Verfahren kein $(2-\varepsilon)$-Approximationsalgorithmus für Vertex Cover ist.

Lösung:

a) Sei G ein Graph und T ein DFS-Spannbaum von G und $I(T)$ wie in der Aufgabe definiert. Da T ein DFS-Spannbaum ist, können zwischen den Blättern von T keine Kanten in G verlaufen, da die T zugrundeliegende Ausführung der Tiefensuche eine solche Kante ansonsten verwendet hätte. Es folgt, dass die Blätter von T ein Independent Set von G sind.

b) Wir zeigen per Induktion über die Knotenzahl n, dass jeder Baum ein Matching besitzt, welches alle internen Knoten von überdeckt. Offensichtlich haben Bäume mit höchstens einem Knoten keine internen Knoten, weswegen die Aussage für $n \leq 1$ erfüllt ist.

Sei nun T ein Baum mit $n+1$ Knoten und sei v ein beliebiges Blatt von T mit Vorgängerknoten w. Wir schneiden nun das Blatt v von T ab und erhalten einen Baum T' auf n Knoten. Nach Induktionsvoraussetzung besitzt T' Matching M', welche alle internen Knoten von T' überdeckt. Somit erhalten wir zwei Fälle:

- M' überdeckt w und somit auch alle internen Knoten von T.
- M' überdeckt w nicht. Dann ist $M' \cup \{vw\}$ ein Matching von T, welches alle internen Knoten von T überdeckt.

Insgesamt hat T also auch ein Matching, welches alle internen Knoten überdeckt.

c) Aus dem ersten Aufgabenteil folgt, dass jede Kante von G mindestens einen Endpunkt in $I(T)$ besitzt. Folglich ist $I(T)$ ein Vertex Cover von G.

Aus dem zweiten Aufgabenteil folgt, dass der DFS-Spannbaum T und somit G ein Matching M besitzt, welches alle Knoten in $I(T)$ überdeckt. Dieses Matching hat Größe $|M| \geq |I(T)|/2$, weil jede Kante von M höchstens 2 Knoten von $I(T)$ überdecken kann. Da jedes Vertex Cover mindestens einen Endpunkt jeder Kante in M wählen muss (siehe Tutoriumsaufgabe 2), haben wir $\text{opt}(G) \geq |M|$ und somit $|I(T)| \leq 2 \cdot \text{opt}(G)$.

Somit ist das Verfahren ein 2-Approximationsalgorithmus für Vertex Cover.

d) Wir betrachten $\mathcal{G} = (C_{2n})_{n\in\mathbb{N}}$, wobei C_k der Kreisgraph auf k Knoten ist.

Unabhängig vom gewählten Startpunkt ist jeder DFS-Spannbaum T von C_{2n} isomorph zu einem gerichteten Pfad der Länge $2n$. Somit hat T genau $2n-1$ interne Knoten, aber offensichtlich gilt $\text{opt}(C_{2n}) = n$. Aus

$$\lim_{n\to\infty} \frac{2n-1}{n} = 2$$

folgt mittels des Cauchy-Kriteriums, dass es für jedes $\varepsilon > 0$ ein n gibt, sodass der Approximationsalgorithmus auf C_{2n} eine Lösung mit Größe $> (2-\varepsilon) \cdot \text{opt}(C_{2n})$ ausgibt.