Aufgabe 1 – *Dominante Menge*

Sei $G = (V, E)$ ein Graph. Eine 'dominante Menge' $W \subseteq V$ ist eine Knotenmenge, sodass für jeden Knoten $v \in V$ gilt, dass entweder v selbst oder ein Nachbar von v in W enthalten ist. In dieser Aufgabe betrachten wir einen randomisierten Algorithmus, der eine dominante Menge findet und dem Algorithmus aus der Vorlesung zum Finden einer stabilen Menge ähnelt.

Wir nehmen an, dass G Minimalgrad mindestens $d > 1$ hat, d.h. jeder Knoten $v \in V$ hat Grad $deg(v) \geq d$. Der Algorithms besteht aus zwei Runden. In der ersten Runde markieren wir jeden Knoten unabhängig von den anderen Knoten mit Wahrscheinlichkeit p (wobei $0 \leq p \leq 1$ gegeben ist). In der zweiten Runde betrachten wir jeden Knoten $v \in V$, wenn weder v noch einer seiner Nachbarn in der ersten Runde markiert wurden, so markieren wir v. Man sieht leicht, dass die Menge der markierten Knoten nach der zweiten Runde eine dominante Menge ist. Wir analysieren die Grösse dieser dominanten Menge im Folgenden.

(a) Sei X die Anzahl der Knoten, die in der ersten Runde markiert werden. Berechnen Sie $\mathbb{E}[X]$.

(b) Sei $v \in V$ ein beliebiger (aber fixer) Knoten. Berechnen Sie die Wahrscheinlichkeit, dass weder v noch einer der Nachbarn von v markiert wurde genau (die Antwort darf von v abhängen). Finden Sie eine obere Schranke für diese Wahrscheinlichkeit, die nur von d und p abhängt (und nicht von v).

(c) Sei Y die Anzahl der Knoten, die in der zweiten Runde markiert werden. Verwenden Sie (b) um eine obere Schranke für $\mathbb{E}[Y]$ zu finden.

(d) Zeigen Sie, dass die erwartete Grösse der dominanten Menge höchstens $n(p + e^{-p(d+1)})$ ist. *Hinweis:* Sie dürfen die Ungleichung $1 - x \leq e^{-x}$ ohne Beweis verwenden.

(e) Zeigen Sie, dass G eine dominante Menge der Grösse höchstens $n \cdot \frac{1+\ln(d+1)}{d+1}$ enthält.

Lösung zu Aufgabe 1 – *Dominante Menge*

Sei $G = (V, E)$ ein Graph mit $V = [n]$. Wir definieren den Wahrscheinlichkeitsraum $\Omega = \{0,1\}^n$. Wobei für $\omega = \omega_1\omega_2 \ldots \omega_n$ gilt das $\Pr[\omega_v = 1] = p$ für alle $v \in V$ unanbängig von allen anderen $\omega_{v'}$. Es ist zu beachten, dass, obwohl unser Algorithmus zweistufig ist, wir den Wahrscheinlichkeitsraum so definieren können, da die zweite Stufe des Algorithmus keinen "neuen" Zufall braucht. In der ersten Runde fügen wir alle Knoten v für die gilt $\omega_v = 1$ in unsere dominante Menge ein. In der zweiten Runde dann alle Knoten w für die gilt $\omega_w = 0$ und $\omega_u = 0$ für alle $u \in N(w)$.

(a) Wir definieren Indikatorzufallsvariablen für jeden Knoten $v \in V$

$$X_v = \begin{cases} 1 & \text{falls } v \text{ in der ersten Runde markiert wurde,} \\ 0 & \text{falls } v \text{ in der ersten Runde nicht markiert wurde.} \end{cases}$$

Dann gilt $X = \sum_{v \in V} X_v$. Des Weiteren gilt

$$\mathbb{E}[X_v] = \Pr[X_v = 1] = \Pr[\omega_v] = p.$$

Durch die Linearität des Erwartungswerts folgt

$$\mathbb{E}[X] = \sum_{v \in V} \mathbb{E}[X_v] = np.$$

(b) Wir definieren das Ereignis

$$F_v = \text{‘weder } v \text{ noch einer seiner Nachbarn wurde in der ersten Runde markiert’}.$$

Für einen beliebigen Knoten w definieren wir das Ereignis

$$E_w = \text{‘}w \text{ wurde in der ersten Runde nicht markiert’}.$$

Es ist leicht zu sehen das gilt

$$\Pr[E_w] = \Pr[\omega_w = 0] = 1 - p.$$

Man beachte das gilt

$$F_v = \bigcap_{w \in \{v\} \cup N(v)} E_w.$$

Wir wissen das die E_w's unanhängig sind, da wir die Knoten in der ersten Stufe unabhängig voneinander zur dominanten Menge hinzufügen. Daraus folgt

$$\Pr[F_v] = \prod_{w \in \{v\} \cup N(v)} \Pr[E_w] = (1-p)^{\deg(v)+1} \leq (1-p)^{d+1},$$

wobei der letzte Schritt folgt weil $1 - p \leq 1$ und somit $(1-p)^{\deg(v)+1}$ monoton fallend in $\deg(v)$ ist.

(c) Wir definieren Indikatorzufallsvariablen für jeden Knoten $v \in V$

$$Y_v = \begin{cases} 1 & \text{falls } v \text{ in der zweiten Runde markiert wurde,} \\ 0 & \text{falls } v \text{ in der zweiten Runde nicht markiert wurde.} \end{cases}$$

Es ist leicht zu sehen, dass nach der Beschreibung des Algorithmus Y_v genau dann 1 ist, wenn F_v eintritt. Aus der Berechnung aus (b) folgt

$$\mathbb{E}[Y_v] = \Pr[Y_v = 1] = \Pr[F_v] \leq (1-p)^{d+1}.$$

Aus der Linearität des Erwartungswerts folgt für $Y = \sum_{v \in V} Y_v$, dass

$$\mathbb{E}[Y] = \sum_{v \in V} \mathbb{E}[Y_v] \leq n(1-p)^{d+1}.$$

(d) Sei Z die Anzahl der Knoten in der dominierenden Menge, die der Algorithmus liefert. Es gilt

$$Z = X + Y.$$

Aus der Linearität des Erwartungswerts folgt:

$$\mathbb{E}[Z] = \mathbb{E}[X] + \mathbb{E}[Y] \leq np + n(1-p)^{d+1}.$$

Durch die in der Aufgabenstellung gegebene Ungleichung erhalten wir $(1-p)^{d+1} \leq e^{-p(d+1)}$. Daraus folgt

$$\mathbb{E}[Z] \leq n\left(p + e^{-p(d+1)}\right).$$

(e) Wir wollen den in (d) erhaltenen Term in Abhängigkeit von p optimieren. Hierfür berechnen wir die erste Ableitung.

$$\frac{\mathrm{d}}{\mathrm{d}p} n\left(p + e^{-p(d+1)}\right) = n\left(1 - (d+1)e^{-p(d+1)}\right)$$

Wir setzen

$$n\left(1 - (d+1)e^{-p(d+1)}\right) \stackrel{!}{=} 0.$$

Auflösen nach p gibt

$$p = \frac{\ln(d+1)}{d+1}.$$

Es ist leicht zu überprüfen das $p \in [0, 1]$.

Man sieht (zum Beispiel über die zweite Ableitung oder anders) leicht, dass es sich bei diesem Wert von p um ein Minimum handelt (und keinen Sattelpunkt/Maximum)

Einsetzen in unsere Formel für $\mathbb{E}[Z]$ gibt

$$\mathbb{E}[Z] \leq n\left(\frac{\ln(d+1)}{d+1} + e^{-\ln(d+1)}\right) = n\left(\frac{1 + \ln(d+1)}{d+1}\right).$$

Dies bedeutet, dass für $p = \frac{\ln(d+1)}{d+1}$ die erwartete Grösse der dominanten Menge höchstens $n\left(\frac{1+\ln(d+1)}{d+1}\right)$ ist. Insbesondere bedeutet dies, dass eine solche dominierende Menge existieren muss (generell gibt es immer Elementarereignisse $\omega_1, \omega_2 \in \Omega$, so dass $X(\omega_1) \geq \mathbb{E}[X]$ und $X(\omega_2) \leq \mathbb{E}[X]$).