HUMBOLDT-UNIVERSITÄT ZU BERLIN

Mathematisch-Naturwissenschaftliche Fakultät II

Institut für Informatik

Die Probabilistische Methode

Proseminar "Das Buch der Beweise"

Dozent: Dr. Wolfgang Kössler

Florian Becker

11. Februar 2013

Inhaltsverzeichnis

1 Einleitung

Neben der klassischen Methode Existenzaussagen direkt durch Konstruktionsbeweise zu zeigen, gibt es auch die sogenannte probabilistische Methode, welche Paul Erdős und Alfred Rényi Mitte des 20. Jahrhunderts innerhalb der Graphentheorie entwickelt haben. Mit dieser ist es oftmals möglich ansonsten sehr aufwendig z.z. Aussagen elegant zu beweisen, indem man nachweist, dass ein bestimmtes Objekt existieren muss, da die Wahrscheinlichkeit dafür echt größer Null ist.

Im folgenden wollen wir nun einige Aussagen zeigen, um diese Methode vorzuführen. Wir geben dabei das gleichnamige Kapitel aus dem Buch *"Das BUCH der Beweise"* [DBdB] von MARTIN AIGNER und GÜNTER M. ZIEGLER wieder.

2 Sätze

2.1 2-färbbare Familien

Sei X eine Obermenge und $\mathcal{F} \subset \mathcal{P}(X)$ eine Familie von Teilmengen mit $|A| = d \geq 2$ für alle $A \in \mathcal{F}$. Wir bezeichnen $\mathcal{F}$ als 2-färbbar, falls eine Färbung $f : X \to \{0, 1\}$ existiert mit

$$\forall A \in \mathcal{F} \exists x, y \in A : \ f(x) \neq f(y).$$

Für $|\mathcal{F}| \geq \binom{2d-1}{d}$ existiert offenbar keine solche Färbung, da für $|X| \geq 2d - 1$ zwangsläufig d Elemente dieselbe Farbe haben.

Satz: *Jede Familie mit höchstens 2^{d-1} d-Mengen ist 2-färbbar.*

Beweis: Sei $|\mathcal{F}| \leq 2^{d-1}$, f eine gleichverteilte zufällige 2-Färbung auf X und bezeichne E_A das Ereignis, dass $A \in \mathcal{F}$ einfarbig ist. Hierfür existieren zwei Möglichkeiten:

$$\Pr[E_A] = 2 \cdot 2^{-d} = 2^{-d+1}$$

Seien nun o.B.d.A die Mengen A_i nicht alle paarweise disjunkt (sonst ist $\mathcal{F}$ trivial 2-färbbar). Damit ist die folgende Ungleichung, in der wir die Vereinigung durch die Summe der Wahrscheinlichkeiten abschätzen, echt:

$$\Pr\left[\bigcup_{A \in \mathcal{F}} E_A\right] < \sum_{A \in \mathcal{F}} \Pr[E_A] = |\mathcal{F}| \cdot \left(\frac{1}{2}\right)^{d-1} \leq 1$$

D.h. die Wahrscheinlichkeit, dass eine einfarbige Menge $A \in \mathcal{F}$ existiert ist echt kleiner Eins. Umgekehrt existiert also eine Färbung von X, sodass $\mathcal{F}$ nur 2-farbige Mengen enthält. $\square$

2.2 Ramsey-Zahlen

Wir bezeichnen mit K_n den vollständigen Graphen auf n Knoten (auch Clique genannt). Sei (a, b) die Eigenschaft, dass bei einer beliebigen 2-Kantenfärbung der Graph K_n eine rote Clique K_a oder eine blaue Clique K_b als Subgraph enthält. Die kleinste solche Zahl n für die (a, b) gilt, bezeichnen wir als Ramsey-Zahl $R(a, b)$.

Satz: *Für die Ramsey-Zahlen gilt mit $k \geq 2$ die untere Schranke: $R(k, k) \geq 2^{\frac{k}{2}}$.*

Beweis: Zunächst betrachten wir die beiden Fälle $k \in \{2, 3\}$. Offensichtlich hat jede Clique mit mindestens 2 Knoten eine rote oder eine blaue Kante, also $R(2, 2) = 2$. Für $k = 3$ betrachten wir den K_5 als Pentagramm, wobei wir die Randkanten rot und die inneren Kanten blau färben. Da dies ein Beispiel für eine dreiecklose Färbung ist, gilt $R(3, 3) > 5$. Wir betrachten nun den K_6 und sei v ein beliebiger Knoten der Clique. Nach dem Schubfachprinzip sind mindestens 3 der 5 Kanten von v einer Farbe zuzuordnen. O.B.d.A. nehmen wir an, dass diese rot sind. Existiert nun zwischen den dadurch implizierten 3 Knoten eine weitere rote Kante, so bildet diese mit v ein rotes Dreieck. Andernfalls sind alle diese Kanten blau und wir haben ein blaues Dreieck. Es ist also $R(3, 3) = 6$.

Sei nun $k \geq 4$, $n < 2^{\frac{k}{2}}$ und wir betrachten eine gleichverteilte zufällige 2-Kantenfärbung von K_n. Bezeichne $A \subset V(K_n)$ eine k-elementige Teilmenge der Knoten und A_R das Ereignis, dass die induzierte Clique $G(A)$ rot ist. Wir berechnen:

$$p_R := \Pr\left[\bigcup_{|A|=k} A_R\right] \leq \sum_{|A|=k} \Pr\left[A_R\right] = \binom{n}{k} \cdot 2^{-\binom{k}{2}}$$

Mit der folgenden groben Abschätzung für den Binomialkoeffizienten

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{1}{k!} \prod_{i=0}^{k-1} n - i \leq \frac{n^k}{k!} \leq \frac{n^k}{2^{k-1}}$$

und dem eingangs gewählten k und n erhalten wir

$$p_R \leq \binom{n}{k} \cdot 2^{-\binom{k}{2}} \leq \frac{n^k}{2^{k-1}} \cdot 2^{-\binom{k}{2}} < 2^{\frac{k^2}{2} - \binom{k}{2} - k + 1} = 2^{-\frac{k}{2}+1} \leq \frac{1}{2}.$$

Analog ergibt für sich blau: $p_B < \frac{1}{2}$. Da nun $p_R + p_B < 1$ ist, folgt für $n < 2^{\frac{k}{2}}$ die Existenz einer Kantenfärbung, sodass K_n weder einen roten noch einen blauen K_k enthält. $\square$

2.3 Chromatische Zahl und Taillenweite

Sei $G = (V, E)$ ein Graph und $f : V \to \mathbb{N}$ eine Färbung der Knoten. f ist eine korrekte Färbung, falls für alle Knoten $v, w \in V$ die Implikation $\{v, w\} \in E \Rightarrow f(v) \neq f(w)$ gilt. Die chromatische Zahl $\chi(G)$ bezeichnet die minmale Anzahl an Farben die notwendig ist für eine korrekte Knotenfärbung des Graphen G.

Auf den ersten Blick liegt die Vermutung nahe, dass eine hohe chromatische Zahl ein Anzeichen dafür ist, dass ein Graph eine große Clique als Untergraphen (und somit kleine Kreise) enthält. Dies ist allerdings im Allgemeinen nicht der Fall, wie der folgende Satz aufzeigt. Hierzu benötigen wir noch den Begriff der Taillenweite $\gamma(G)$ eines Graphen, welche die Länge eines kürzesten Kreises in G bezeichnet und die Stabilitätszahl $\alpha(G)$, die die Kardinalität einer größten stabilen Menge bezeichnet.

Satz: *Für jedes $k \geq 2$ gibt es einen Graphen G mit chromatischer Zahl $\chi(G) > k$ und Taillenweite $\gamma(G) > k$.*

Beweis: Sei $\mathcal{G}_{n,p}$ ein Wahrscheinlichkeitsraum über Graphen mit n Knoten und einer Kantenwahrscheinlichkeit $p \in [0, 1]$ und $G \in \mathcal{G}_{n,p}$ ein Zufallsgraph. Da eine korrekte Färbung die Knotenmenge eines Graphen vollständig partitioniert in stabile Mengen, ergibt sich die Ungleichung $\chi\alpha \geq n$. Indem wir nun also zeigen, dass die Stabilitätszahl α klein ist, muss die Färbungszahl χ entsprechend groß sein.

Sei $2 \leq r \leq n$. Wir schätzen die Wahrscheinlichkeit ab, dass die Stabilitätszahl eine gewisse Mindestgröße hat. Dazu nutzen wir die bekannte Ungleichung $1 + x \leq e^x, \forall x \in \mathbb{R}$:

$$\begin{aligned} \Pr[\alpha \geq r] &\leq \binom{n}{r}(1-p)^{\binom{r}{2}} \leq n^r(1-p)^{\binom{r}{2}} = \left(n(1-p)^{\frac{r-1}{2}}\right)^r \\ &\leq \left(n\exp(-p)^{\frac{r-1}{2}}\right)^r = \left(n\exp\left(-\frac{p(r-1)}{2}\right)\right)^r \end{aligned}$$

Wir setzen nun $r := \left\lceil \frac{n}{2k} \right\rceil$ und $p := n^{-\frac{k}{k+1}}$. Da $n^{\frac{1}{k+1}} \in \omega(\log n)$, gilt ab genügend großem n die Ungleichung:

$$\begin{aligned} n^{\frac{1}{k+1}} &\geq 6k\log n \mid \cdot n^{-1} \\ n^{-\frac{k}{k+1}} = p &\geq 6k\frac{\log n}{n} \mid \cdot r \\ pr &\geq 3\log n \end{aligned}$$

Damit folgt:

$$\begin{aligned} n\exp\left(-\frac{p(r-1)}{2}\right) &= n\exp\left(-\frac{pr}{2}\right)\exp\left(\frac{p}{2}\right) \leq ne^{-\frac{3}{2}\log n}e^{\frac{1}{2}} \\ &= n^{-\frac{1}{2}}e^{\frac{1}{2}} = \left(\frac{e}{n}\right)^{\frac{1}{2}} \longrightarrow 0, \text{ für } n \to \infty \end{aligned}$$

Da $\sqrt{\frac{e}{n}}$ eine Nullfolge ist und $r = \left\lceil \frac{n}{2k} \right\rceil$ unbeschränkt mit n wächst, folgt ebenfalls:

$$\Pr\left[\alpha \geq r\right] \leq \left(\frac{e}{n}\right)^{\frac{r}{2}} \longrightarrow 0, \text{ für } n \to \infty$$

D.h. ab einem gewissen n_1 gilt für alle $n \geq n_1$die Schranke $\Pr\left[\alpha \geq \frac{n}{2k}\right] < \frac{1}{2}$. Damit haben wir den ersten notwendigen Teil unseres Beweises gezeigt.

Als nächstes widmen wir uns der Taillenweite. Damit diese eine gewisse Mindestgröße erreicht, wollen wir sicherstellen, dass es in dem betrachteten Graphen nicht zu viele Kreise der Größe $\leq k$ gibt. Sei hierzu X die Zufallsvariable welche diese Kreise zählt und seien $X_C : \Omega \to \{0, 1\}$ Indikatorvariablen, welche für den Kandidaten $C \subset V(G)$ bestimmen, ob es sich um einen Kreis in G handelt. Offensichtlich gilt mit Länge i für den Erwartungswert: $\mathbb{E}\left[X_C\right] = \Pr\left[C \text{ ist Kreis in } G\right] = p^i$. Weiterhin gibt es als mögliche Knotenfolgen $n \cdot (n-1) \cdot \ldots \cdot (n-i+1)$ Kombinationen, wobei sich dabei jeweils $2i$ Variationen auf denselben Kreis beziehen aufgrund unterschiedlicher Startknoten und Richtungen. Mit der Linearität des Erwartungswertes ergibt sich somit als mittlere Anzahl an Kreisen der Länge $\leq k$:

$$\mathbb{E}\left[X\right] = \sum_{i=3}^{k} \frac{n!}{(n-i)! \cdot 2i} p^i \leq \frac{1}{2} \sum_{i=3}^{k} (np)^i \leq \frac{1}{2}(k-2)(np)^k$$

Wobei wir genutzt haben, dass $np = n^{\frac{1}{k+1}} \geq 1$ ist. Indem wir nun die Markov-Ungleichung auf X anwenden, erhalten wir:

$$\begin{aligned} \Pr\left[X \geq \frac{n}{2}\right] &\leq \frac{2 \cdot \mathbb{E}\left[X\right]}{n} \leq (k-2)\frac{(np)^k}{n} = (k-2)\frac{n^{\frac{k}{k+1}}}{n} \\ &= (k-2)n^{-\frac{1}{k+1}} \longrightarrow 0, \text{ für } n \to \infty \end{aligned}$$

Es existiert also ein n_2, sodass für alle $n \geq n_2$ die Schranke $\Pr\left[X \geq \frac{n}{2}\right] < \frac{1}{2}$ erfüllt ist. Jetzt können wir endlich die eigentliche Aussage zeigen.

Sei nun $n \geq \max\{n_1, n_2\}$. Dann ist die Wahrscheinlichkeit, dass $\alpha < \frac{n}{2k}$ und $X < \frac{n}{2}$ gilt, echt größer Null. Sei H solch ein Graph. Um die geforderte Taillenweite zu erreichen, entfernen wir nun aus jedem Kreis der Länge $\leq k$ einen Knoten (also höchstens $\left\lceil \frac{n}{2} \right\rceil - 1$) und bezeichnen diesen Graphen als F. Damit gilt nun $\gamma(F) > k$ und offensichtlich ist $\alpha(F) \leq \alpha(H)$. Für die chromatische Zahl folgt nun mit dem schon überlegten Zusammenhang $\chi\alpha \geq n$:

$$\chi(F) \geq \frac{|V(F)|}{\alpha(F)} \geq \frac{n}{2\alpha(H)} > \frac{n}{2\frac{n}{2k}} = k$$

Wir haben damit nun die Existenz eines Graphen mit den geforderten Eigenschaften bewiesen. $\square$

2.4 Kreuzungslemma

Als letztes Beispiel für die probabilistische Methode betrachten wir die sogenannte Kreuzungszahl eines Graphen. Sei hierzu G ein Graph mit n Knoten und m Kanten. Die minimale Anzahl an Kantenkreuzungen die für eine Zeichnung bzw. eine Einbettung von G in den $\mathbb{R}^2$ nötig sind, bezeichnen wir als seine Kreuzungszahl $\mathrm{cr}(G)$. Ist die Kreuzungszahl gleich Null, so ist der Graph planar. Die Eulerformel (auch: Polyedersatz) gibt eine Bedingung für die Kantenanzahl planarer Graphen: $m \leq 3n-6$. Anhand dieser können wir leicht eine erste untere Schranke für die Kreuzungszahl ermitteln.

Liege hierzu eine bzgl. $\mathrm{cr}(G)$ minimale Einbettung von G vor. Wir konstruieren den Graphen H indem wir nun alle Kreuzungspunkte zu Knoten machen. H hat dann $n + \mathrm{cr}(G)$ Knoten und $m + 2\mathrm{cr}(G)$ Kanten, da jeder neue Kreuzungsknoten Grad 4 hat. Da H selbst nun ein planarer Graph ist, gilt die Eulerformel und wir können für beliebige Graphen G ableiten:

$$\begin{aligned} & m + 2\mathrm{cr}(G) \leq 3(n + \mathrm{cr}(G)) - 6 = 3n + 3\mathrm{cr}(G) - 6 \\ \Longrightarrow \quad & m - 3n + 6 \leq \mathrm{cr}(G) \end{aligned}$$

Allerdings können wir eine bessere untere Schranke angeben, wenn die Kantenanzahl asymptotisch echt schneller wächst als die Knotenanzahl.

Satz: *Sei G ein einfacher Graph mit n Ecken und m Kanten, wobei $m \geq 4n$ gelten soll. Dann ist*

$$cr(G) > \frac{1}{64}\frac{m^3}{n^2}.$$

Beweis: Sei eine minimale Einbettung von G gegeben und $p \in [0,1]$. Wir erzeugen einen Subgraphen G_p, indem wir jeden Knoten von G mit Wahrscheinlichkeit p wählen und die dadurch induzierten Kanten nehmen. Bezeichne weiterhin n_p, m_p und X_p Zufallsvariablen, welche die Anzahl der Knoten, Kanten und Kreuzungen in G_p zählen. Nach unserer vorherigen Überlegung gilt:

$$\begin{aligned} & X_p \geq \mathrm{cr}(G_p) \geq m_p - 3n_p + 6 \\ \Longleftrightarrow \quad & X_p - m_p + 3n_p \geq 6 > 0 \\ \Longrightarrow \quad & \mathbb{E}\left[X_p - m_p + 3n_p\right] > 0 \end{aligned}$$

Wir nutzen wieder die Linearität des Erwartungswertes aus und berechnen die einzelnen Summanden. Es gilt offensichtlich $\mathbb{E}\left[n_p\right] = pn$ und $\mathbb{E}\left[m_p\right] = p^2 m$. Für die Anzahl der Kreuzungen müssen jeweils alle 4 beteiligten Knoten aus der Einbettung von G ausgewählt worden sein, es ergibt sich also $\mathbb{E}\left[X_p\right] = p^4\mathrm{cr}(G)$. Damit folgt:

$$\begin{aligned} & 0 < \mathbb{E}\left[X_p\right] - \mathbb{E}\left[m_p\right] + 3\mathbb{E}\left[n_p\right] = p^4\mathrm{cr}(G) - p^2 m + 3pn \\ \Longleftrightarrow \quad & \mathrm{cr}(G) > \frac{p^2 m - 3pn}{p^4} = \frac{m}{p^2} - \frac{3n}{p^3} \end{aligned}$$

Nun setzen wir $p = \frac{4n}{m} \leq 1$ und erhalten

$$\operatorname{cr}(G) > \frac{m}{p^2} - \frac{3n}{p^3} = \frac{m^3}{16n^2} - \frac{3nm^3}{64n^3} = \frac{1}{64}\left(\frac{4m^3}{n^2} - \frac{3m^3}{n^2}\right) = \frac{1}{64}\frac{m^3}{n^2},$$

womit wir die z.z. Aussage bewiesen haben. $\square$

Zum Schluss wollen wir noch beide Schranken vergleichen. Wächst nun die Kantenanzahl asymptotisch schneller als die Knotenanzahl, also $m \in \Theta(n^c)$ mit einem $c > 1$, so ist unsere erste Schranke asymptotisch beschränkt durch n^c, während das Kreuzungslemma ein asymptotisches Wachstum von $n^{3c-2} > n^c$ aufweist. Das Kreuzungslemma liefert uns also tatsächlich eine asymptotisch schärfere untere Schranke, falls die Kantenanzahl superlinear mit der Knotenanzahl wächst.

Literatur

[DBdB] Martin Aigner, Günter M. Ziegler.
'35. Die Probabilistische Methode'
In: *Das BUCH der Beweise*. Auflage 2, 2003, pp. 257-265.
ISBN: 3-540-40185-7