Fakulta Matematiky, Fyziky a Informatiky
Univerzita Komenského, Bratislava

Štúdium niektorých vlastností náhodne indukovaných podgrafov n-rozmernej hyperkocky

Bc. Lenka Trojaková

2011

Katedra Informatiky
Fakulta Matematiky, Fyziky a Informatiky
Univerzita Komenského, Bratislava

Štúdium niektorých vlastností náhodne indukovaných podgrafov n-rozmernej hyperkocky

b301b51b-7a59-4dd4-a058-909d4bad9409

Bc. Lenka Trojaková

Diplomová práca

9.2.1 Informatika

vedúci práce: doc. DNDr. Eduard Toman, CSc.

Bratislava, 2011

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Lenka Trojaková
Študijný program: informatika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: 9.2.1. informatika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský

Názov: Štúdium niektorých vlastností náhodne indukovaných podgrafov n-rozmernej hyperkocky

Cieľ: Zvládnutie základných pravdepodobnostných metód pri štúdiu istých náhodných premenných. Získať zručnosť pri určovaní odhadov ich veľkosti.

Vedúci: doc. RNDr. Eduard Toman, CSc.

Dátum zadania: 25.02.2010

Dátum schválenia: 18.02.2011

prof. RNDr. Branislav Rovan, PhD.
garant študijného programu

študent

vedúci

Prehlasujem, že som svoju prácu vypracovala samostatne
s využitím mojich poznatkov, skúseností a vedomostí môjho vedúceho
a s použitím uvedenej literatúry.

Bratislava, máj 2011 *Lenka Trojaková*

Chcela by som poďakovať môjmu vedúcemu
doc. RNDr. Eduardovi Tomanovi, CSc.
za jeho čas, cenné rady pri riešení problémov
a pripomienky k tejto diplomovej práci.

Taktiež ďakujem za podporu mojim kamarátom a rodine,
hlavne mojej sestre Erike za to,
že si napriek všetkému našla čas a chuť
na prečítanie tejto práce a debaty o nej.

Abstrakt

Boolovské funkcie majú veľa praktických využití. Ako príklad spomenieme ich súvis s kryptografiou, symetrickou šifrou, či navrhovaním obvodov a čipov pre výpočtovú techniku.
Problém zjednodušenia boolovských funkcií sa dá previesť na problém vrcholového pokrytia podgrafu hyperkocky menšími hyperkockami. Keďže tento problém je NP-úplný, jeden z možných prístupov k jeho riešeniu je zmenšovanie množiny vrcholov, ktoré treba pokryť. Tu sa dostávame k pojmom regulárny vrchol a jadrová podkocka. Nakoľko sa zaoberáme náhodnými funkciami, venujeme sa v práci aj základným charakteristikám niektorých náhodných premenných súvisiacich s jadrovými podkockami. Zaoberáme sa strednou hodnotou počtu jadrových podkociek v náhodnom grafe a odhadujeme disperziu počtu takýchto podkociek.
Zaujímavá je aj otázka počtu iredundantných disjunktívnych normálnych foriem pre náhodnú booleovskú funkciu. Odpoveď na ňu úzko súvisí s pokrývaním náhodného podgrafu hyperkocky maximálnymi podkockami. V práci poskytujeme horný a dolný odhad počtu týchto foriem, ktoré získavame na základe výsledkov niekoľkých autorov, ktorí sa venovali tematike náhodne indukovaných podgrafov hyperkocky.

Kľúčové slová: Náhodné grafy, prahové funkcie, jadrová podkocka, regulárny vrchol, iredundantná disjunktívna normálna forma.

Abstract

Usage of boolean functions can be found in many practical areas. Cryptography, symmetric cypher, designing of circuits and microchips for information technologies are just a few of them.
Problem of simplifying boolean function is reducable to vertex cover problem of hypercube by smaller cubes. This problem is NP-complete and one of the possible approaches to simplify the problem is decreasing the number of vertices to be covered. That way we get to concept of kernel cubes and regular vertices. While we work with random functions, we devote space for basic characteristics of some random variables connected with kernel cubes. We inspect the expected value of number of kernel cubes in random graph. We also estimate the variance of this random variable.
Finding the number of irredundant disjunctive normal forms of random boolean function is essential for satisfiability of boolean function. Both these problems are strongly connected with covering of random subgraf of hypercube by smaller hypercubes. In this thesis we provide upper and lower bound of number of these forms. These results are based on results of several mathematicians who have already spent time researching the field of randomly induced subgraphs of hypercube.

Key words: Random graph, threshold function, kernel cube, regular vertex, iredundant disjunctive normal form.

Obsah

Úvod

Teória náhodných grafov sa začala rozvíjať v 20. storočí. Prvýkrát ju v roku 1959 rozpracovali Paul Erdös a Alfrér Rényi v práci s názvom "On Random Graphs" (O náhodných grafoch) a nezávisle na nich Edgar Gilbert vo svojej práci "Random Graphs" (Náhodné grafy). Odvtedy bolo v tejto teórii dosiajhnutých mnoho zaujímavých výsledkov.

My sa v našej práci zameriavame na vlastnosti náhodných podgrafov hyperkocky a ich súvis s booleovskými funkciami. Zaoberáme sa charakteristickými vlastnosťami niektorých náhodných premenných súvisiacich s pokrývaním grafu podgrafmi danej triedy a počtom iredundantných disjunktných normálnych foriem pre náhodnú booleovskú funkciu.

V práci nadväzujeme na výsledky viacerých autorov. Distribúciou dimenzie maximálnych podkociek sa vo svojej práci zaoberali autori E. Toman, Ľ. Baník a M. Stanek ([5]). Jadrovým podkockám, strednej hodnote ich počtu a asymptotickému odhadu ich počtu sa venoval E. Toman v ďalšej zo svojich prác ([6]). Vlastnostiam náhodných booleovských funkcií sa venuje vo svojich dvoch prácach M. Škoviera ([7],[8]).

V prvej kapitole práce uvádzame základné definície, pojmy a niektoré vety, ktoré budeme vyžívať v celej práci. Je rozdelená na dve podčasti. Prvá obsahuje pojmy z teórie grafov, druhá sa zaoberá teóriou pravdepodobnosti.

V druhej kapitole sa venujeme prahovým funkciám vlastností náhodných podgrafov nejakého grafu. Ozrejmíme, čo vlastne pojem prahovej funkcie predstavuje. Venujeme sa metóde zisťovania prahovej funkcie. V závere kapitoly pomocou tejto metódy určíme prahovú funkciu pre vlastnosť definovanú ako obsahovanie maximálnej podkocky v náhodnom podgrafe hyperkocky.

Tretia kapitola pojednáva o vrcholovom pokrytí podgrafu hyperkocky, s ktorým úzko súvisia jadrové podkocky a regulárne vrcholy. Zaoberáme sa strednou hodnotou počtu jadrových podkociek a ohraničujeme disperziu počtu takýchto podkociek.

Posledná, štvrtá kapitola, rozoberá problematiku počtu iredundantných disjunktívnych normálnych foriem náhodných booleovských funkcií. V prvej časti prezentujeme okrem iného algoritmus na zostrojenie iredundantných disjunktívnych normálnych foriem a na jeho základe odvodíme a dokážeme dolný odhad pre počet týchto foriem. V druhej časti uvádzame horný odhad pre počet iredundantných disjunktívnych normálnych foriem.

V práci používame kombinatoricko-pravdepodobnostné metódy, výpočet stredných hodnôt a disperzie pre rôzne typy náhodných premenných, na ktorých získanie vytvárame pomocné veličiny. Markovova, Čebyševova nerovnosť a rôzne kombinatorické odhady sú často skryté za výpočtami v tejto práci.

Kapitola 1

Základné definície a pojmy

V nasledujúcej kapitole uvedieme kvôli presnosti, lepšej čitateľnosti a jednoznačnosti pojmy, ktoré sú použité v tejto práci. Sú to prevažne pojmy z teórie grafov a teórie pravdepodobnosti. Takmer všetky pojmy v tejto aj nasledujúcej kapitole sú prebraté z použitej lieratúty.

1.1 Niektoré pojmy z teórie grafov

Definícia 1.1. Graf G je usporiadaná dvojica (V, E), kde V je nejaká neprázdna množina a E je množina dvojprvkových podmnožín množiny V. Prvky množiny V nazývame vrcholy grafu G a prvky množiny E hrany grafu G.

Definícia 1.2. Hyperkocka Q_n je graf (V, E), ktorý má nasledujúce vlastnosti:

1. množina vrcholov V je množina všetkých binárnych vektorov dĺžky n;
2. hrana v grafe je medzi každými dvoma vrcholmi, ktorých vektory sa líšia práve na jednom mieste.

Z druhej vlastnosti hyperkocky ľahko vidieť, že každý vrchol je stupňa n. Keďže hyperkocka Q_n má 2^n vrcholov, mohutnosť množiny hrán je $|E| = n \cdot 2^{n-1}$.

Definícia 1.3. Hyperkocka G rádu k sa nazýva podkockou hyperkocky Q_n, ak G je indukovaným podgrafom Q_n.

Definícia 1.4. Nech H je podgraf Q_n. Hyperkocka $G \subseteq H$ je maximálnou podkockou rádu k, ak G je podkocka rádu k a neexistuje hyperkocka G' taká, že $G \subset G' \subseteq H$.

Lema 1.5. Počet podkociek rádu k hyperkocky Q_n je $\binom{n}{k} \cdot 2^{n-k}$.

Dôkaz. Podkocka rádu k je určená k „voľnými" bitmi vo vektoroch zodpovedajúcich jej vrcholom, ostatných $n - k$ bitov je „pevných". Je $\binom{n}{k}$ spôsobov, ako vybrať k rôznych pozícií a 2^{n-k} rôznych binárnych vektorov pre „pevné" bity. $\square$

1.2 Pojmy z teórie pravdepodobnosti

Definícia 1.6. Nech Ω je množina elementárnych udalostí, pričom $\Omega \neq \emptyset$ a nech $\mathcal{S} \subseteq 2^{\Omega}$. Neprázdny systém $\mathcal{S}$ podmnožín množiny elementárnych udalostí nazývame pole náhodných udalostí (a ozn. σ-pole), ak spĺňa nasledovné podmienky:

1. $\Omega \in \mathcal{S}$;
2. ak $A \in \mathcal{S}$, potom $A^C \in \mathcal{S}$, kde $A^C = \Omega - A$;
3. ak $A_i \in \mathcal{S}$ pre $i = 1, 2, \ldots$, potom $\cup_{i=1}^{\infty} A_i \in \mathcal{S}$.

Definícia 1.7. Nech $(\Omega, \mathcal{S})$ je σ-pole. Prvky množiny Ω nazveme elementárne výsledky a podmnožiny Ω patriace do systému $\mathcal{S}$ nazveme udalosti.

Definícia 1.8. Pravdepodobnostná miera na σ-poli udalostí $(\Omega, \mathcal{S})$ je zobrazenie $P : \mathcal{S} \to \mathbb{R}$ spĺňajúce nasledovné podmienky:

1. pre všetky $A \in \mathcal{S}$ platí $0 \leq P(A) \leq 1$;
2. $P(\Omega) = 1$, $P(\emptyset) = 0$;
3. ak $(A_i)_{i=1}^{n}$ je postupnosť disjunktných udalostí, tak $P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$.

Definícia 1.9. Nech $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ je množina elemetárnych rovnako pravdepodobných vzájomne sa vylučujúcich udalostí. Nech m z týchto udalostí, $m \leq n$, má za následok nastatie udalosti A, potom pravdepodobnosť udalosti A definujeme ako podiel

$$P(A) = \frac{m}{n}.$$

Definícia 1.10. Nech $(\Omega, \mathcal{S})$ je σ-pole udalostí a nech P je pravdepodobnostná miera na $(\Omega, \mathcal{S})$. Potom trojicu $(\Omega, \mathcal{S}, P)$ nazveme pravdepodobnostný priestor.

Definícia 1.11. Nech (Ω, S, P) je pravdepodobnostný priestor. Budeme hovoriť, že funkcia $X : \Omega \to \mathbb{R}$ je náhodná premenná, ak pre každé $x \in \mathbb{R}$ platí

$$\{\omega \in \Omega : X(\omega) < x\} \in \mathcal{S}.$$

Definícia 1.12. Náhodnú premennú X na pravdepodobnostnom priestore $(\Omega, \mathcal{S}, P)$ nazývame diskrétna, ak jej obor hodnôt $X(\Omega) \subset \mathbb{R}$ je spočítateľná množina.

Definícia 1.13. Nech X je diskrétna náhodná premenná, ktorá nadobúda hodnoty $(x_i)_{i \in I}$ s nenulovou pravdepodobnosťou. Potom stredná hodnota náhodnej premennej X je definovaná nasledovne

$$E(X) = \sum_{i \in I} x_i P[X = x_i].$$

(V prípade, že suma $\sum_{i \in I} x_i P[X = x_i]$ neexistuje, hovoríme, že X nemá strednú hodnotu.)

Definícia 1.14. Strednú hodnotu funkcie $g(X) = (X - E(X))^k$ nazývame centrálnym momentom k-teho rádu a zapisujeme

$$\mu_k = E[(X - E(X))^k] = \sum_{i \in I} (x_i - E(X))^k p_i.$$

Ak $k = 2$, potom $\mu_2 = E(X - E(X))^2 = D(X)$ nazývame aj disperziou (rozptylom).

Veta 1.15. Nech X a Y sú diskrétne náhodné premenné, ktoré majú konečnú strednú hodnotu. Nech a, b sú reálne čísla. Potom aj diskrétna náhodná premenná $aX + bY$ má konečnú strednú hodnotu a platí

$$E(aX + bY) = aE(X) + bE(Y).$$

Veta 1.16. Nech X je diskrétna náhodná premenná s konečnou strednou hodnotou aj disperziou. Potom platí:

- $D(X) = E(X^2) - (E(X))^2$;
- $D(aX + b) = a^2 D(X)$ pre všetky $a, b \in \mathbb{R}$.

Veta 1.17 (Markovova nerovnosť). Nech X je nezáporná náhodná premenná. Potom pre každé $\varepsilon > 0$ platí

$$P[X \geq \varepsilon] \leq \frac{E(X)}{\varepsilon}.$$

Veta 1.18 (Čebyševova nerovnosť). Nech X je diskrétna náhodná premenná s konečnou disperziou (t.j. aj s konečnou strednou hodnotou). Potom pre každé $\delta > 0$ platí

$$P([X - E(X)] \geq \delta) \leq \frac{D(X)}{\delta^2}.$$

Kapitola 2

Prahová funkcia

V tejto kapitole budeme pracovať s pravdepodobnostným priestorom, v ktorom $\mathcal{G}_n$ označuje množinu všetkých podgrafov n-rozmernej hyperkocky a každá hrana sa v grafe vyskytuje s pravdepodobnosťou p závislou od n.

Budeme sledovať, ako sa s rastom pravdepodobnosti p mení pravdepodobnosť, že vzniknutý podgraf H grafu G má danú vlastnosť $\mathcal{P}$, čo budeme označovať ako $H \in \mathcal{P}$. Konkrétnejšie, budeme sa snažiť určiť, pre aké p podgraf H danú vlastnosť má, pre aké p podgraf H vlastnosť $\mathcal{P}$ takmer určite nemá, a hlavne sa budeme snažiť určiť kritickú pravdepodobnosť, pre akú sa nadobúdanie danej vlastnosti mení.

Táto kritická hodnota sa nazýva prahová funkcia. Budeme ju označovať $t(n)$. Ak vezmeme funkciu asymptoticky menšiu ako $t(n)$, tak takmer žiadny graf danú vlastnosť nemá. Pre pravdepodobnosť p asymptoticky väčšiu ako prahová funkcia takmer každý graf má uvažovanú vlastnosť $\mathcal{P}$.

Existencia prahovej funkcie pre danú vlastnosť nie je samozrejmá; existujú vlastnosti, ktoré prahovú funkciu nemajú. Takouto vlastnosťou je napríklad existencia Hamiltonovskej kružnice v náhodne indukovanom podgrafe ([1]).

Definícia 2.1. Reálnu funkciu $t(n)$, ktorá nadobúda len kladné hodnoty, nazývame prahovou funkciou vlastnosti $\mathcal{P}$, ak pre každú pravdepodobnosť $p(n)$ a každý graf $G \in \mathcal{G}_n$ platí

$$\lim_{n\to\infty} P[G \in \mathcal{P}] = \begin{cases} 0, \text{ ak } p/t \to 0 \text{ pre } n \to \infty \\ 1, \text{ ak } p/t \to \infty \text{ pre } n \to \infty. \end{cases}$$

Nech má vlastnosť $\mathcal{P}$ prahovú funkciu $t(n)$. Potom aj každý kladný násobok ct funkcie t je tiež prahovou funkcicou pre $\mathcal{P}$. Prahová funkcia spĺňajúca vyššie uvedené vlastnosti je teda určená jednoznačne až na multiplikatívnu konštantu. Vezmime si už spomínanú vlastnosť existencie Hamiltonovskej kružnice a uvažujme pravdepodobnosť $p = (1+\varepsilon)\ln n.n^{-1}$. Je známe, že pre $\varepsilon > 0$ obsahuje graf Hamiltonovský cyklus, no pre ε záporné to už zaručiť nevieme.

Metóda určovania prahovej funkcie

Všeobecná metóda na určenie prahovej funkcie je v [1] popísaná nasledovne:

Uvažujme *nezápornú* náhodnú premennú X definovanú na $\mathcal{G}_n$ a vlastnosť $\mathcal{P}$, kde

$$\mathcal{P} = \{G \mid X(G) > 0\}.$$

Potrebujeme ukázať, že $\mathcal{P}$ má prahovú funkciu t. Dôkaz sa skladá z dvoch častí:

1. dokážeme, že takmer žiadny graf $G \in \mathcal{G}_n$ nemá vlastnosť $\mathcal{P}$, ak pravdepodobnosť p je asymptoticky menšia ako t, teda $p/t \to 0$ pre $n \to \infty$.

2. dokážeme, že takmer každý graf $G \in \mathcal{G}_n$ má vlastnosť $\mathcal{P}$, ak pravdepodobnosť p je asymptoticky väčšia ako $t(n)$, čiže $p/t \to \infty$ pre $n \to \infty$.

Dôkaz prvej časti: Ak X je celočíselná funkcia, použitím Markovovej nerovnosti s $\varepsilon = 1$ nájdeme horné ohraničenie pre $P[X \geq 1]$ vyjadrené pomocou strednej hodnoty. Ak $\lim_{n\to\infty} E(X) = 0$, potom $X(G)$ môže byť kladná a teda rovná najmenej 1 len pre malú množinu grafov $G \in \mathcal{G}_n$, čiže väčšina grafov vlastnosť $\mathcal{P}$ nemá.

Výhoda tohto postupu je v tom, že je oveľa jednoduchšie počítať strednú hodnotu ako samotné pravdepodobnosti, nemusíme sa zaoberať nezávislosťou či nezlúčiteľnosťou udalostí. Keďže stredná hodnota je lineárna, môžeme strednú hodnotu náhodnej premennej vypočítať ako súčet stredných hodnôt indikátorových náhodných premenných.

Dôkaz druhej časti je komplikovanejší. Aby sme ukázali, že $P[X > 0]$ je dostatočne veľká, nestačí ohraničiť $E(X)$ zdola, lebo náhodná premenná X

nie je ohraničená zhora. Môže sa stať, že $E(X)$ je veľká len vďaka tomu, že náhodná premenná X je veľká len pre zopár grafov G a teda môže byť stále rovná nule pre väčšinu grafov $G \in \mathcal{G}_n$.
Aby sme ukázali, že $P[X > 0] \to 1$, musíme ukázať, že takáto situácia nenastane, to znamená, že X sa príliš nelíši od svojej strednej hodnoty. Na tento účel môžeme použiť Čebyševovu nerovnosť, pričom však musíme určiť disperziu náhodnej premennej. Určovanie disperzie je však vo veľa prípadoch obtiažne, lebo pre jej výpočet potrebujeme určiť strednú hodnotu druhej mocniny náhodnej premennej X. Táto komplikácia sa často dá obísť tým, že nájdeme horný odhad disperzie.

Po určení disperzie použijeme Čebyševovu nerovnosť nasledovne:

Lema 2.2. Ak $E(X) > 0$ pre $n \to \infty$ a $D(X)/E(X)^2 \to 0$ pre $n \to \infty$, potom $X(G) > 0$ pre takmer všetky $G \in \mathcal{G}_n$.

Dôkaz. Pre ľubovoľný graf G, pre ktorý $X(G) = 0$, platí $|X(G) - E(X)| = E(X)$. Potom z Čebyševovej nerovnosti, kde položíme $\delta = E(X)$, vyplýva, že

$$P[X = 0] \leq P[X \leq 0] + P[X \geq 2E(X)] = P[|X - E(X)| \geq E(X)] \leq \frac{D(X)}{E(X)^2}.$$

Pritom $D(X)/E(X)^2 \to 0$ pre $n \to \infty$ a preto $X(G) > 0$ pre skoro všetky grafy $G \in \mathcal{G}_n$. $\square$

Dôsledok 2.3. Ak $E(X) > 0$ pre $n \to \infty$ a $E(X^2)/E(X)^2 \to_{n \to \infty} 1$, potom $X(G) > 0$ pre takmer všetky $G \in \mathcal{G}_n$.

V práci [3] sme použitím tejto metódy dokázali nasledovnú vetu:

Veta 2.4. Nech $\mathcal{P}$ označuje vlastnosť, že náhodný podgraf hyperkocky obsahuje maximálnu podkocku rádu k. Potom

$$t(n) = \left(\sqrt[k \cdot 2^{k-1}]{2^n} \cdot \sqrt[2^{k-1}]{n} \right)^{-1}$$

je prahová funkcia pre $\mathcal{P}$.

Dôkaz. Podkocky rádu k danej hyperkocky Q_n označíme $K_1, K_2, \ldots, K_r$. Z lemy 1.5 vieme, že ich počet $r = \binom{n}{k} \cdot 2^{n-k}$.

Nech $\mathcal{I}_j$ je náhodná premenná definovaná takto:

$$\mathcal{I}_j(G) = \begin{cases} 1, \text{ ak graf } G \text{ obsahuje danú podkocku } K_j \\ 0 \quad \text{v opačnom prípade.} \end{cases}$$

Ľahko vypočítame, že

$$E(\mathcal{I}_j) = P[\mathcal{I}_j = 1] = p^{k \cdot 2^{k-1}} \cdot \left(1 - p^{(k+2)2^{k-1}}\right)^{n-k}.$$

Výraz $p^{k \cdot 2^{k-1}}$ je rovný pravdepodobnosti, že v grafe G bude podkocka K_j, teda že graf G bude obsahovať všetkých $k \cdot 2^{k-1}$ hrán k-rozmernej podkocky. Zvyšok výrazu tvorí pravdepodobnosť, že v G nie je žiadna z podkociek rádu $k+1$ obsahujúcich K_j.

Nech $\mathcal{X}_k(G)$ je náhodná premenná, ktorá je rovná počtu maximálnych podkociek rádu k v grafe G. Potom

$$\mathcal{X}_k(G) = \sum_{1 \le j \le r} \mathcal{I}_j.$$

Preto

$$\begin{aligned} E(\mathcal{X}_k(G)) &= E\left(\sum_{1 \le j \le r} \mathcal{I}_j\right) = r \cdot E(\mathcal{I}_1) \\ &= \binom{n}{k} 2^{n-k} \cdot p^{k \cdot 2^{k-1}} \cdot \left(1 - p^{(k+2)2^{k-1}}\right)^{n-k}. \end{aligned}$$

Teraz dokážeme, že pre pravdepodobnosti asymptoticky menšie ako $t(n)$ nemá takmer žiadny graf vlastnosť $\mathcal{P}$.

Nech $p(n)/t(n) = \gamma(n)$. Predpokladáme, že $p(n)$ je asymptoticky menšie ako $t(n)$. čiže $\lim_{n \to \infty} \gamma(n) = 0$.

Z vety 1.17 vyplýva, že stačí ukázať rovnosť $\lim_{n\to\infty} E(\mathcal{X}_k) = 0$. Rátajme

$$\begin{aligned}
E(\mathcal{X}_k) &= \binom{n}{k} 2^{n-k} \cdot p^{k\cdot 2^{k-1}} \cdot \left(1 - p^{(k+2)2^{k-1}}\right)^{n-k} \\
&= \binom{n}{k} 2^{n-k} \cdot \left(\gamma(n) \cdot \left(\sqrt[k\cdot 2^{k-1}]{2^n} \cdot \sqrt[2^{k-1}]{n}\right)^{-1}\right)^{k\cdot 2^{k-1}} \\
&\quad \cdot \left(1 - \left(\gamma(n) \cdot \left(\sqrt[k\cdot 2^{k-1}]{2^n} \cdot \sqrt[2^{k-1}]{n}\right)^{-1}\right)^{(k+2)2^{k-1}}\right)^{n-k} \\
&= \binom{n}{k} 2^{n-k} \cdot \frac{\gamma(n)^{k\cdot 2^{k-1}}}{2^n \cdot n^k} \cdot \left(1 - \frac{\gamma(n)^{(k+2)2^{k-1}}}{(2^n)^{\frac{k+2}{k}} \cdot n^{k+2}}\right)^{n-k} \\
&= \binom{n}{k} \cdot \frac{\gamma(n)^{k\cdot 2^{k-1}}}{2^k \cdot n^k} \cdot \left(1 - \frac{\gamma(n)^{(k+2)2^{k-1}}}{(2^n)^{\frac{k+2}{k}} \cdot n^{k+2}}\right)^{n-k}.
\end{aligned}$$

Nakoľko

$$\lim_{n\to\infty} \frac{\gamma(n)^{k\cdot 2^{k-1}}}{2^k} \cdot \left(1 - \frac{\gamma(n)^{(k+2)2^{k-1}}}{(2^n)^{\frac{k+2}{k}} \cdot n^{k+2}}\right)^{n-k} = 0,$$

keďže ide o limitu súčinu funkcií, z ktorých prvá má limitu 0 a druhá je ohraničená, ľahko na základe tej istej vety o limite súčinu funkcií dostávame, že

$$\begin{aligned}
\lim_{n\to\infty} E(\mathcal{X}_k) &= \lim_{n\to\infty} \binom{n}{k} \cdot \frac{\gamma(n)^{k\cdot 2^{k-1}}}{2^k \cdot n^k} \cdot \left(1 - \frac{\gamma(n)^{(k+2)2^{k-1}}}{(2^n)^{\frac{k+2}{k}} \cdot n^{k+2}}\right)^{n-k} \\
&= 0.
\end{aligned}$$

Ešte potrebujeme ukázať, že pre pravdepodobnosti väčšie ako $t(n)$ majú takmer všetky grafy vlastnosť $\mathcal{P}$.

Zaveďme si novú náhodnú premennú $\mathcal{X}_{i,j}$ takto:

$$\mathcal{X}_{i,j} = \begin{cases} 1, \text{ ak } K_i \text{ a } K_j \text{ sú maximálne } k\text{-rozmerné podkocky v } G, \\ 0 \ \text{ inak.} \end{cases}$$

Pre $i = 1, 2, \ldots, r$ bude A_i označovať udalosť, že kocka K_i je maximálna k-rozmerná podkocka grafu G. Pod *okolím* podkocky K budeme rozumieť všetky vrcholy hyperkocky Q_n, z ktorých vedie hrana do nejakého vrchola podkocky K (kocka K je súčasťou svojho okolia). *Dobrá dvojica* podkociek K_i, K_j je taká neusporiadaná dvojica, že prienik okolí kociek K_i a K_j je prázdny. Dvojicu podkociek, ktorá nie je dobrá, budeme nazývať *zlou dvojicou.*

Z každého vrchola hyperkocky G môže vychádzať najviac n hrán, z ktorých najviac k patrí podkocke K a zvyšné patria G, ale nepatria K (*trčia* z K). Koncové vrcholy hrán trčiacich z jednotlivých vrcholov kocky K musia byť navzájom rôzne. Dokážeme toto tvrdenie sporom. Ak by vrchol $v \notin K$ mal v K aspoň dvoch susedov x a y, tak medzi x a y existuje cesta dĺžky 2. Pritom x a y nemôžu byť spojené hranou (bipartitný graf Q_n neobsahuje trojuholníky), čiže najkratšia možná cesta medzi nimi má dĺžku 2. Lenže všetky cesty minimálnej dĺžky medzi dvomi vrcholmi z K patria do K, preto aj vrchol v musí patriť do K, čo je spor s jeho voľbou.

Z úvahy v predošlom odseku vyplýva, že veľkosť okolia podkocky rádu k je $(n-k+1)2^k$.

Spomeňme si teraz na pojem zlej dvojice podkociek a určime horný odhac ich počtu. K danej k-rozmernej podkocke K grafu Q_n spočítame, s najviac koľkými kockami L môže tvoriť zlú dvojicu. Okolie kocky L musí mať s okolím kocky K spoločný aspoň jeden vrchol; kocka L je určená svojou telesovou uhlopriečkou, čiže druhým vrcholom tejto uhlopriečky.

Keď máme danú podkocku K, tak počet kociek, s ktorými tvorí zlú dvojicu, je nanajvýš $(n-k+1)2^k \cdot \binom{n}{k}$, nakoľko $(n-k+1)2^k$ je počet možností, ako vybrať prvý vrchol uhlopriečky podkocky L v okolí podkocky K a $\binom{n}{k}$ je počet možností, ako vybrať druhý vrchol — z n súradníc vyberieme k, ktoré budeme meniť.

Na základe predchádzajúceho odstavca a faktu, že počet možností pre výber podkocky K je $r = 2^{n-k} \cdot \binom{n}{k}$ vieme ohraničiť počet zlých dvojíc podkociek m takto:

$$m \leq 2^{n-k} \cdot \binom{n}{k}(n-k+1) \cdot 2^k \cdot \binom{n}{k} = 2^n \cdot (n-k+1) \cdot \binom{n}{k}^2. \qquad (2.1)$$

Poďme teraz odhadnúť $E(\mathcal{X}_{i,j})$. Označme $\mathcal{D}$ množinu všetkých dobrých dvojíc podkociek. Keďže $E(\mathcal{X}_{i,j}) = P(A_i \cap A_j)$, nastávajú dve možnosti:

- ak (K_i, K_j) je dobrá dvojica podkociek, tak $E(\mathcal{X}_{i,j}) = P(A_i) \cdot P(A_j)$
- inak $E(\mathcal{X}_{i,j}) = P(A_i \cap A_j) \leq 1$

Ľahko vidieť, že

$$E(\mathcal{X}_k^2) = E\left(\sum_{1\leq i,j\leq r} \mathcal{X}_{i,j}\right) = \sum_{1\leq i,j\leq r} E(\mathcal{X}_{i,j}),$$

a teda

$$\begin{aligned} E(\mathcal{X}_k^2) &= \sum_{(K_i,K_j)\in\mathcal{D}} E(\mathcal{X}_{i,j}) + \sum_{(K_i,K_j)\notin\mathcal{D}} E(\mathcal{X}_{i,j}) \\ &\leq \sum_{(K_i,K_j)\in\mathcal{D}} P(A_i) \cdot P(A_j) + 2^n \cdot (n-k+1)\binom{n}{k}^2 \cdot 1 \end{aligned}$$

Vieme, že $P(A_i) = P(A_j) = P[\mathcal{X}_1 = 1] = p^{k\cdot 2^{k-1}} \cdot (1 - p^{(k+2)2^{k-1}})^{n-k}$ pre dobrú dvojicu kociek. Počet všetkých dobrých dvojíc môžeme zhora ohraničiť počtom všetkých dvojíc k-rozmerných podkociek, teda $(\binom{n}{k} \cdot 2^{n-k})^2$. Dostávame, že

$$\begin{aligned} E(\mathcal{X}_k^2) &\leq \binom{n}{k}^2 \cdot (2^{n-k})^2 \cdot \left(p^{k\cdot 2^{k-1}} \cdot (1 - p^{(k+2)2^{k-1}})^{n-k}\right)^2 \\ &+ 2^n \cdot (n-k+1)\binom{n}{k}^2. \end{aligned}$$

Dôkaz zakončíme použitím 2.3, čím ukážeme, že $X(G) > 0$ pre skoro všetky grafy z $\mathcal{G}_n$.

$$\begin{aligned} \lim_{n\to\infty} \frac{E(\mathcal{X}_k^2)}{E(\mathcal{X}_k)^2} &\leq \lim_{n\to\infty} \frac{\binom{n}{k}^2 \cdot (2^{n-k})^2 \cdot \left(p^{k\cdot 2^{k-1}} \cdot (1 - p^{(k+2)2^{k-1}})^{n-k}\right)^2}{\binom{n}{k}^2 \cdot (2^{n-k})^2 \cdot \left(p^{k\cdot 2^{k-1}} \cdot (1 - p^{(k+2)2^{k-1}})^{n-k}\right)^2} \\ &+ \lim_{n\to\infty} \frac{2^n \cdot (n-k+1)\binom{n}{k}^2}{\binom{n}{k}^2 \cdot (2^{n-k})^2 \cdot \left(p^{k\cdot 2^{k-1}} \cdot (1 - p^{(k+2)2^{k-1}})^{n-k}\right)^2} \\ &\leq 1 + \lim_{n\to\infty} \frac{2^k \cdot (n-k+1)}{2^{n-k} \cdot \left(p^{k\cdot 2^{k-1}} \cdot (1 - p^{(k+2)2^{k-1}})^{n-k}\right)^2}. \end{aligned}$$

Nech $p(n)/t(n) = \gamma(n)$. Tentokrát predpokladáme, že $p(n)$ je asymptoticky väčšie ako $t(n)$. čiže $\lim_{n\to\infty} \gamma(n) = \infty$. Potom

$$\begin{aligned}
\lim\nolimits_{n\to\infty} & \frac{(n-k+1)}{2^{n-2k}\cdot\left(p^{k\cdot 2^{k-1}}\cdot(1-p^{(k+2)2^{k-1}})^{n-k}\right)^2} \\
= & \lim_{n\to\infty} \frac{n}{2^{n-2k}\cdot\left((\gamma(n)\cdot t(n))^{k\cdot 2^{k-1}}\cdot(1-(\gamma(n)\cdot t(n))^{(k+2)2^{k-1}})^{n-k}\right)^2} \\
= & \lim_{n\to\infty} \frac{n\cdot 2^n}{(n^k)^2\cdot(\gamma^{k\cdot 2^{k-1}})^2\cdot\left(\left(1-\gamma^{(k+2)2^{k-1}}\cdot(2^{-n})^{\frac{k+2}{k}}\right)^{n-k}\right)^2\cdot 2^{n-k}} \\
= & \lim_{n\to\infty} \frac{n\cdot 2^n}{(n^k)^2\cdot(\gamma^{k\cdot 2^{k-1}})^2\cdot e^0\cdot 2^{n-k}} \\
= & \ 0,
\end{aligned}$$

pre $p(n)$ asymptoticky menšie ako $q(n) = (n-k)^{\frac{-1}{(k+2)2^{k-1}}}$ (pre takéto p bude vo výraze vystupovat „iba“ e^0). Čiže

$$\lim_{n\to\infty} E(\mathcal{X}_k^2)/E(\mathcal{X}_k)^2 = 1,$$

čo sme chceli dokázať.

$\square$

Poznámka 2.1. Existenciu hraničnej pravdepodobnosti $q(n)$ sme už vopred očakávali na základe toho, že pre $p \to 1$ začína byť podgraf hyperkocky príliš hustý a teda k-rozmerné podkocky nie sú maximálne, lebo sú súčasťou väčších kociek.

Kapitola 3

Vrcholové pokrytie hyperkocky

Definícia 3.1. Vrcholové pokrytie grafu G je množina $S \subseteq V(G)$ taká, že každá hrana grafu je incidentná s aspoň jedným vrcholom z množiny S.

Je jednoduché nájsť nejaké vrcholové pokrytie daného grafu. Stačí zobrať $S = V(G)$. Zaujímavejší a hlavne zložitejší je problém vrcholového pokrytia s danou maximálnou veľkosťou, minimálneho vrcholového pokrytia resp. minimálneho pokrytia grafmi danej triedy. Keďže napriek dlhoročnému úsiliu nepoznáme efektívny algoritmus na ich riešenie, mnohí sa snažia zlepšiť známe algoritmy.

Problém vrcholového pokrytia má veľa praktických aplikácií v teórii grafov, ale aj v iných oblastiach. Ako príklad možno uviesť zjednodušovanie Boolovských funkcií z triedy disjunktných normálnych formúl, ktoré sa dá zredukovať na problém vrcholového pokrytia n-rozmernej hyperkocky menšími kockami.

Pozrime sa bližšie na redukciu spomínanú v predchádzajúcom odseku. Majme boolovskú funkciu $f\colon \{0,1\}^3 \to \{0,1\}$ danú nasledovnou tabuľkou:

a	1	1	1	1	0	0	0	0
b	1	1	0	0	1	1	0	0
c	1	0	1	0	1	0	1	0
$f(a,b,c)$	1	1	1	0	1	1	0	0

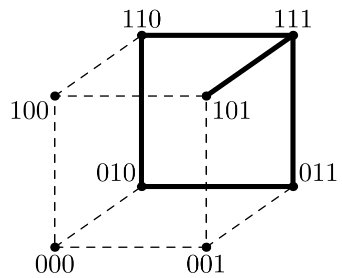
Z tabuľky vidieť, že funkcia f sa dá zapísať ako

$$f(a,b,c) = abc \vee ab\overline{c} \vee a\overline{b}c \vee \overline{a}bc \vee \overline{a}b\overline{c},$$

kde využívame označenie $xyz = x \wedge y \wedge z$.

Zostrojíme graf k funkcii f, ktorý bude podgrafom nejakej hyperkocky a ukážeme, ako súvisí minimálne pokrytie tohto grafu maximálnymi podkockami so zjednodušením funkcie f.

Definičný obor funkcie f tvoria usporiadané trojice – binárne vektory. Vektoru (x, y, z) zodpovedá binárny reťazec xyz, resp. vrchol hyperkocky označený týmto reťazcom. Vezmime si podgraf hyperkocky indukovaný vrcholmi, ktoré zodpovedajú trojiciam (x, y, z) takým, že $f(x, y, z) = 1$. Pre funkciu definovanú vyššie uvedenou tabuľkou dostaneme graf G zobrazený na obrázku 3.1.



Obr. 3.1: Graf zodpovedajúci funkcii f.

Pokrytím grafu G maximálnymi podkockami je množina

$$\mathcal{S} = \{\{010, 011, 110, 111\}, \{101, 111\}\}.$$

Je ľahké určiť, ktoré zložky binárnych vektorov označujúcich vrcholy danej podkocky v pokrytí sa zachovávajú, a ktoré sa menia. Ak máme n-rozmernú hyperkocku (funkcia f má n premenných) a v pokrytí máme k-rozmernú podkocku, tak sa zachováva $n-k$ zložiek vektorov pre vrcholy a teda zvyšných k zložiek sa mení.

Ako toto súvisí so zjednodušovaním zápisu danej funkcie?
Máme graf G, ktorého vrcholy sú označené reťazcami $x_1 x_2 \ldots x_n$. Pre každú maximálnu podkocku v pokrytí grafu G – množine $\mathcal{S}$ – vytvoríme zodpovedajúcu konjunkciu niekoľkých premenných podľa nasledujúcich pravidiel:

- ak je zložka x_i rôzna aspoň pre dva vrcholy danej podkocky, tak zodpovedajúcu premennú zo zápisu vylúčime;

- ak je zložka $x_i = 1$ pre všetky vrcholy danej podkocky, tak zodpovedajúca premenná bude vo výslednej konjunkcii literálov v tvare x_i;
- ak je zložka $x_i = 0$ pre všetky vrcholy danej podkocky, tak zodpovedajúca premenná bude vo výslednej konjunkcii literálov negovaná, teda v tvare $\overline{x_i}$.

Keď máme všetky konjunkcie $C_1, C_2, \ldots, C_l$, ostáva len skonštruovať výslednú formulu pridaním disjunkcií a dostávame $C_1 \vee C_2 \vee \cdots \vee C_l$.

Pre našu funkciu f a množinu $\mathcal{S}$, ktorá je pokrytím zodpovedajúceho grafu G dostávame nasledovné konjunkcie:

- podkocke $\{010, 011, 110, 111\}$ zodpovedá konjunkcia b;
- podkocke $\{101, 111\}$ zodpovedá konjunkcia ac.

Výsledná zjednodušená formula pre funkciu f je teda $b \vee ac$.

Čím je funkcia jednoduchšia, tým ľahšie sa určí, čí je splniteľná, čo je tiež jeden z NP-úplných problémov. Ak by sme vedeli pokrývať indukované podgrafy hyperkocky maximálnymi podkockami, tak vieme zjednodušovať boolovské funkcie. Hoci proces zjednodušovania pomocou pokrytí vyzerá zložito v porovnaní s použitím, napríklad, Karnaughových máp, pre väčšie množstvo premenných sa nestáva neprehľadným.

Regulárny vrchol, jadrová množina

Pre ďalšie pokračovanie práce uvedieme alternatívnu definíciu vrcholového pokrytia grafu.

Definícia 3.2. Majme graf $G = (V, E)$. Nech $\mathcal{S} = \{S_1, S_2, \ldots, S_m\}$ je množina niektorých podmnožín množiny V. Každú množinu $\mathcal{S}' \subseteq \mathcal{S}$ takú, že

$$\bigcup_{A \in \mathcal{S}'} A = V,$$

nazývame *pokrytím* grafu G.
Jadrová množina grafu G je taká množina $S_k \subseteq V$, že existuje vrchol $v \in S_k$, pričom pre všetky $j \neq k$ platí: $v \notin S_j$. Je to teda množina, ktorá obsahuje

vrchol nepokrytý žiadnou inou množinou a preto musí táto množina určite patriť do pokrytia grafu G. Takúto množinu nazývame tiež jadrový interval. Nech $V_i = \{j | v_i \in S_j\}$. Ak existujú rôzne $r, s \in \{1, \ldots, n\}$ také, že $V_r \subseteq V_s$, tak každá množina v $\mathcal{S}$ obsahujúca vrchol v_r obsahuje aj vrchol v_s. Takýto vrchol v_s sa nazýva *regulárny vrchol* grafu G a je to vrchol, ktorý je pokrytý, ak je pokrytý vrchol v_r.

V nasledujúcom odstavci budeme používať označenie $V \to V'$ pre situáciu, že do množiny V priradíme množinu V'. Ak graf obsahuje jadrovú množinu S_k, tak môžeme zjednodušiť problém vrcholového pokrytia daného grafu takto:

$$V \to V \setminus S_k,\ \mathcal{S} \to \{S_1 \setminus S_k, S_2 \setminus S_k, \ldots, S_{k-1} \setminus S_k, S_{k+1} \setminus S_k, \ldots, S_m \setminus S_k, \}.$$

Ak graf obsahuje regulárny vrchol v_s, opäť môžeme problém vrcholového pokrytia zjednodušiť, a to nasledovne:

$$V \to V \setminus \{v_s\},\ \mathcal{S} \to \{S_1 \setminus \{v_s\}, S_2 \setminus \{v_s\}, \ldots, S_m \setminus \{v_s\}\}.$$

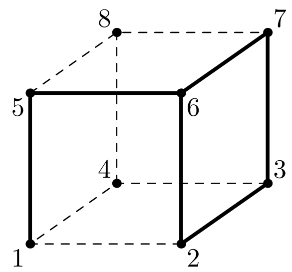
V obidvoch prípadoch sme zmenšili množinu vrcholov aspoň o 1. Čím viac regulárnych vrcholov a jadrových množín obsahuje daný graf, tým viac vieme pomocou nich zjednodušiť problém vrcholového pokrytia tohto grafu.

3.1 Jadrové podkocky

V niekoľkých ďalších častiach sa budeme venovať náhodne indukovaným podgrafom n-rozmernej hyperkocky, pričom vrchol sa v grafe vyskytuje s pravdepodobnosťou p závislou od n.

Definícia 3.3. Maximálna podkocka K hyperkocky Q_n sa nazýva *jadrová podkocka*, ak existuje vrchol $v \in V(K)$ taký, že pre všetky maximálne podkocky $J \neq K$ obsiahnuté v Q_n platí: $v \notin J$. Množina všetkých jadrových podkociek grafu G sa nazýva *jadro grafu G*.

Na nasledujúcom obrázku je znázornený príklad podgrafu G 3-rozmernej hyperkocky, ktorý obsahuje dve jadrové podkocky $K_1 = (1, 5)$ a $K_2 = (2, 3, 7, 6)$. Podkocka $L = (6, 5)$ nie je jadrová.



Obr. 3.2: Náhodný graf G s jadrovými podkockami K_1 a K_2

Nech C_k je náhodná premenná označujúca počet k-rozmerných jadrových podkociek, jej strednú hodnotu označíme $E(C_k)$. Ďalej nech C je náhodná premenná označujúca počet všetkých jadrových pokociek a $E(C)$ nech je stredná hodnota tejto náhodnej premennej.

Teraz uvedieme hodnoty $E(C_k)$ a $E(C)$, ktorých skúmaním sa zaoberá práca [6]. Začneme nasledujúcou lemou.

Lema 3.4. Nech K je podkocka Q_n rádu k a nech $G \in \mathcal{G}_n$. Potom

$$P[\text{K je jadrová}|K \subseteq G] = 1 - (1 - (1 - p)^{n-k})^{2^k}.$$

Dôkaz. Označme hranu, ktorá má spoločný vrchol s podkockou K, ale nepatrí do K, ako *okolitú.*
Hodnota $1 - p$ predstavuje pravdepodobnosť, že v K sa nevyskytuje nejaká hrana h. Po umocnení tejto hodnoty na $(n - k)$ dostávame pravdepodobnosť, že pre nejaký vrchol podkocky K sa žiadna z okolitých hrán tohto vrchola nenachádza v G.
Odčítaním pravdepodobnosti $(1 - p)^{(n-k)}$ od 1 dostávame pravdepodobnosť, že aspoň jedna okolitá hrana konkrétneho vrchola sa nachádza v grafe G. Umocnenie na 2^k upravuje túto pravdepodobnosť pre všetky vrcholy podkocky K.
Odčítaním hodnoty $(1 - (1 - p)^{n-k})^{2^k}$ od 1 tak dostávame pravdepodobnosť, že G neobsahuje žiadnu z okolitých hrán, čiže K je jadrová. □

Veta 3.5. Nech $p \in \langle 0, 1\rangle$ je pravdepodobosť, že indukovaný podgraf hyperkocky Q_n obsahuje ľubovoľný vrchol. Potom platí

$$E(C_k) = \binom{n}{k} 2^{n-k} p^{2^k} \left(1 - \left(1 - (1 - p)^{n-k}\right)^{2^k}\right). \tag{3.1}$$

Dôkaz. Zaveďme indikátor $\eta_k(G)$ nasledovne:

$$\eta_K(G) = \begin{cases} 1, & \text{ak podkocka } K \text{ je jadrová v } G, \\ 0 & \text{v opačnom prípade.} \end{cases}$$

Platí nasledujúci vzťah:

$$E(C_k) \;=\; E\Big(\sum_{K\in G} \eta_K(G)\Big) = \sum_{K\in G} P\Big(\eta_K(G) = 1\Big).$$

Vieme, že $P(\eta_K(G) = 1)$ je rovnaká pre všetky podkocky K a je rovná výrazu $p^{2^k} \cdot \Big(1 - \big(1 - (1-p)^{n-k}\big)^{2^k}\Big)$.
Hodnota p^{2^k} je pravdepodobnosť označujúca existenciu podkocky K v G a na základe lemy 3.4 vieme, že $\Big(1 - \big(1 - (1-p)^{n-k}\big)^{2^k}\Big)$ predstavuje pravdepodobnosť, že K je jadrová v G.

Spojením týchto skutočností a lemy 1.5 dostávame vyššie uvedený výsledok.

□

Dôsledok 3.6. Z predchádzajúcej vety ľahko vidieť, že platí:

$$E(C) = \sum_{k=1}^{n} \binom{n}{k} 2^{n-k} p^{2^k} \left(1 - \big(1 - (1-p)^{n-k}\big)^{2^k}\right).$$

Pre ďalšie skúmanie vlastností náhodnej premennej C_k bude užitočné poznať disperziu tejto náhodnej premennej. Označme $D(C_k)$ disperziu náhodnej premennej C_k. V nasledujúcej vete uvádzame horný odhad pre $D(C_k)$ a v dôkaze spôsob, ako sme sa k nemu dopracovali.

Veta 3.7. Nech $P_K = p^{2^k} \left(1 - \big(1 - (1-p)^{n-k}\big)^{2^k}\right)$ je pravdepodobnosť, že daná k-rozmerná kocka K obsiahnutá v G je jadrová. Potom

$$D(C_k) \leq \binom{n}{k}^2 2^{n-k} \cdot (1 - P_K^2).$$

Dôkaz. Na výpočet $D(C_k)$ potrebujeme dve základné hodnoty. Jednou je $E(C_k)$, ktorú už máme, druhou je $E(C_k^2)$, ktorú sa pokúsime ohraničiť.

Zvolme indikátor $\eta_k(G)$ rovnako ako v dôkaze vety 3.1. Potom platí

$$\begin{aligned} E(C_k^2) &= E\Big(\sum_{K\in G}\eta_K^2(G)\Big) = \sum_{(K,L)\in G} E\Big(\eta_K(G)\cdot\eta_L(G)\Big) \\ &= \sum_{(K,L)\in G} P\Big(\eta_K(G)\cdot\eta_L(G)=1\Big). \end{aligned}$$

$$P\big(\eta_K(G)\cdot\eta_L(G)=1\big)\begin{cases} P\big(\eta_K(G)=1\big)\cdot P\big(\eta_L(G)=1\big), \text{ ak } K\cap L=\emptyset \\ \leq 1 \quad \text{v opačnom prípade.}\end{cases}$$

Dvojicu k-rozmerných kociek (K,L) nazvime *zlou* vtedy, keď $K\cap L\neq\emptyset$. Ostatné dvojice (K,L) nazveme *dobré* a ich počet označme d_k. Nech l_k označuje počet zlých dvojíc (K,L) v G. Počet spôsobov, ako vybrať kocku K je $\binom{n}{k}\cdot 2^{n-k}$. Nech prienikom K a L je nejaká j-rozmerná kocka. Kocku L potom môžeme dovyberať $\sum_{j=0}^{k}\binom{k}{j}\binom{n-k}{k-j}=\binom{n}{k}$ spôsobmi. Najskôr vyberieme j súradníc, ktoré budú mať K a L spoločné (čím zabezpečíme j-rozmerný prienik) a potom dovyberáme zo zvyšných $n-k$ súradníc ostávajúcich $k-j$, aby sme získali celú kocku L. Preto

$$l_k = 2^{n-k}\binom{n}{k}^2.$$

Z tohto výsledku a lemy 1.5 dostávame, že

$$d_k = \binom{n}{k}^2 2^{n-k}\Big(2^{n-k}-1\Big),$$

a teda

$$E(C_k^2)\leq d_k\cdot P_K^2 + 2^{n-k}\binom{n}{k}^2.$$

Spojením tohto vzťahu, vzťahu 3.1 a vzorca pre výpočet disprezie náhodnej premennej po niekoľkých úpravách dostávame požadovaný výsledok, konkrétne

$$\begin{aligned} D(C_k) &\leq \binom{n}{k}^2 2^{n-k}\cdot(1-P_K^2) \\ &\leq \binom{n}{k}^2 2^{n-k}\cdot\left(1-p^{k\cdot 2^k}\left(1-\big(1-(1-p)^{n-k}\big)^{2^k}\right)^2\right). \end{aligned}$$

□

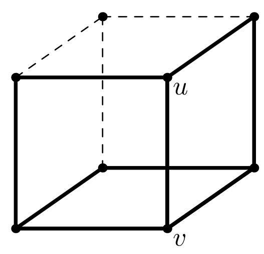
Odhad strednej hodnoty a disperzie náhodnej premennej C_k sme určovali hlavne za účelom zistenia prahovej funkcie pre vlastnosť obsahovania jadrovej podkocky. Toto úsilie však neviedlo k želanému výsledku. Zisťovanie prahových funkcií však nie je jediný spôsob využitia disperzie a strednej hodnoty náhodnej premennej. Preto uvádzame odhady týchto hodnôt aj so spôsobom ich získania.

3.2 Regulárne vrcholy

Regulárne vrcholy sú ďalšou štruktúrou, ktorú môžeme skúmať aj v náhodne indukovaných podgrafoch hyperkocky. Pre korektnosť uveďme najskôr definíciu.

Definícia 3.8. Nech G je náhodne indukovaný podgraf Q_n. Vrchol $v \in V(G)$ sa nazýva *regulárny*, ak existuje vrchol $u \neq v$ taký, že pre všetky maximálne podkocky K grafu G platí: $u \in K \Rightarrow v \in K$.

Na nasledujúcom obrázku môžeme vidieť príklad regulárneho vrchola. Ak je pokrytý vrchol u nejakou maximálnou podkockou, tak je pokrytý aj vrchol v.



Obr. 3.3: Náhodný podgraf hyperkocky s regulárnym vrcholom v.

Nech R_G je náhodná premenná označujúca počet regulárnych vrcholov v grafe G. Podľa [4] platí

$$R_G < \varrho^n,$$

kde $1{,}51 \leq \varrho < 2$ je konštanta závislá od p.

Vlastnosti regulárnych vrcholov a ohraničenie ich počtu využijeme v nasledujúcej kapitole.

Kapitola 4

Iredundatné disjunktné normálne formy náhodných booleovských funkcií

V tejto kapitole sa budeme venovať iredundantným disjunktným normálnym formám (DNF) náhodných booleovských funkcií, ktoré, ako sme ukázali v tretej kapitole, úzko súvisia s pokrývaním náhodného podgrafu hyperkocky maximálnymi podkockami. Zameriame sa hlavne na dolný a horný odhad počtu takýchto formúl.

Definícia 4.1. Pokrytie U množiny N_f maximálnymi podkockami budeme nazývať iredundantné, ak žiadna vlastná podmnožina U nepokrýva N_f. Iredundantná DNF je tá, ktorá zodpovedá takémuto pokrytiu.

4.1 Dolný odhad počtu iredundantných DNF

Na zistenie dolného odhadu použijeme náhodné booleovské funkcie n premenných, ktoré s pravdepodobnosťou idúcou k 1 pre n idúce do nekonečna majú nasledujúce vlastnosti:

1. $f(x_1, \ldots, x_n)$ nadobúda hodnotu 1 aspoň na $p \cdot 2^n - \varphi(n)\sqrt{p \cdot 2^n}$ vrcholoch Q_n – n-rozmernej hyperkocky, kde $\frac{1}{\varphi(n)} = o(1)$.

2. Rozmer maximálnych podkociek f nie je väčší ako $\mu = \lg \log_b n + 2$, kde $b = \frac{1}{p}$.
3. Počet jadrových podkociek $c(f) < n^{(1+o(1)) \lg \log_b n}(2(1-p))^n$.
4. Počet regulárnych bodov $r(f)$ nie je väčší ako ρ^n, kde $1.51 \leq \rho < 2$ je konštanta závislá od p.

Vyššie uvedené výsledky sú prebraté z prác [4], [7], [8] a využijeme ich pri dôkaze nasledujúcej lemy.

Lema 4.2. Nech f je náhodná booleovská funkcia n premenných, $M_f \subseteq N_f$ je množina vrcholov n-rozmernej hyperkocky, ktoré nie sú pokryté jadrovými intervalmi a nie sú regulárnymi bodmi. Označme $|M_f|$ ako počet prvkov množiny M_f. Potom

$$|M_f| \geq 2^n(1 - o(1)).$$

Dôkaz. Vieme, že rozmer jadrových intervalov náhodnej booleovskej funkcie nie je väčší ako μ a nie je ich viac ako $c(f)$. Jadrové intervaly teda pokrývajú najviac $c(f) \cdot 2^\mu$ bodov. Okrem toho $r(f) < \rho^n$. Teda

$$\begin{aligned}
|M_f| &\geq p \cdot 2^n - \varphi(n)\sqrt{p \cdot 2^n} - \rho^n - c(f) \cdot 2^\mu \\
&\geq p \cdot 2^n - \varphi(n)\sqrt{p \cdot 2^n} - \rho^n \\
&- n^{(1+o(1)) \lg \log_{1/p} n}(2(1-p))^n \cdot 2^{\lg \log_b n + 2} \\
&= 2^n\Big(p - o(1) - 4(1-p)^n \log_b n \cdot n^{(1+o(1)) \lg \log_{1/p} n}\Big) \\
&= 2^n\big(p - o(1)\big) = 2^n(1 - o(1)).
\end{aligned}$$

□

Nasledujúcu lemu spolu s definíciami využijeme pri určovaní dolného odhadu počtu iredundantných DNF.

Definícia 4.3. Blokový (n, d) kód (E, D) pozostáva z kódovacej funkcie $E : 2^n \to 2^d$, dekódovacej funkcie $D : 2^d \to 2^n$, pričom funkcia $E \circ D$ je funkcia identity a funkcie E, D spracovávajú text po blokoch dĺžky n, resp. d.

Definícia 4.4. Grupovým kódom $G_{n,d}$ s parametrom d budeme nazývať ľubovoľný blokový (n, d) kód, ktorého všetky kódové slová tvoria aditívnu podgrupu.

Lema 4.5. Nech M je ľubovoľná podmnožina n-rozmernej hyperkocky Q_n, $|M|$ je počet vrcholov v M a nech v Q_n existuje grupový kód $G_{n,d}$ so vzdialenosťou d. Potom existuje podmnožina $M' \subseteq M$ taká, že vzdielenosť medzi ľubovoľnými vrcholmi M' je aspoň d a platí

$$|M'| \geq |M| \frac{|G_{n,d}|}{2^n}.$$

Dôkaz. Rozložíme Q_n na invariantné triedy podľa podgrupy $G_{n,d}$.

$$Q_n = G_{n,d} \cup G_{n,d}^1 \cup \cdots \cup G_{n,d}^{\frac{2^n}{g}},$$

kde $g = |G_{n,d}|$.

Každá z $G_{n,d}^i$ je kód so vzdialenosťou d. Tento kód indukuje rozklad

$$M = (M \cap G_{n,d}) \cup (M \cap G_{n,d}^1) \cup \cdots \cup (M \cap G_{n,d}^{\frac{2^n}{g}-1}).$$

Aspoň jedna z týchto $\frac{2^n}{g}$ množín obsahuje aspoň $|M| \cdot \frac{g}{2^n}$ bodov. Týmto je lema dokázaná. □

4.1.1 Algoritmus na zostrojenie iredundantných DNF

Opíšeme proces zostrojenia iredundatných d. n. f. pre náhodnú booleovskú funkciu. Budeme konštruovať zodpovedajúce pokrytia množiny N_f. Algoritmus pozostáva zo štyroch hlavných krokov.

1. krok Začleníme do pokrytia jadrové podkocky.

Nech $M_f \subseteq N_f$ (ako aj predtým) je množina vrcholov, ktoré nie sú pokryté jadrovými podkockami a nie sú regulárnymi bodmi. Z práce [9] je známe, že v n-rozmernej hyperkocke Q_n existuje grupový kód so vzdialenosťou $2r+1$ s počtom prvkov $g \geq \frac{2^n}{(2n)^r}$. Vytvoríme množinu $\overline{M}_f \subseteq M_f$ tak, že bude obsahovať body, ktorých vzájomná vzdialenosť je aspoň $2\mu+1$, kde hodnotu μ sme uviedli v úvode kapitoly. Na základe lemy 4.5 môžeme predpokladať, že počet prvkov $|\overline{M}_f|$ vyhovuje nerovnosti

$$|\overline{M}_f| \geq \frac{|M_f|}{(2n)^\mu}.$$

2. krok Pokrývame každý vrchol množiny $\overline{M}_f$ niektorou podkockou, ktorá obsahuje daný vrchol.

3. krok Pokrývame každý zo zostávajúcich nepokrytých bodov podkockou, ktorá neobsahuje žiaden bod z množiny $\overline{M}_f$.

Dokážeme, že tento krok možno vždy uskutočniť. V opačnom prípade existuje bod $\tilde{\beta}$ taký, že ľubovoľný maximálny interval prechádzajúci cez tento bod aj cez body $\overline{M}_f$. Dôsledkom toho, že vzdialenosť bodov z $\overline{M}_f$ je aspoň $2\mu + 1$ a rozmer maximálnych intervalov nie je väčší ako μ, posledná podmienka implikuje, že všetky intervaly, ktoré prechádzajú cez $\tilde{\beta}$, prechádzajú cez jeden a ten istý vrchol $\tilde{\alpha} \in \overline{M}_f$. To však znamená, že $\tilde{\alpha}$ je regulárny bod, čo protirečí konštrukcii množiny $\overline{M}_f$.

4. krok Vynechávame z pokrytia podkocky, ktoré sme zaradili do pokrytia v prvých dvoch krokoch a ktoré sa nachádzajú v zjednotení ostatných, až do získania iredundantného pokrytia.

Pri 4. kroku nebude žiadna z podkociek zaradených v druhom kroku vyhodený, pretože každá z nich obsahuje bod z $\overline{M}_f$ pokrytý len jednou podkockou. Takýmto spôsobom získame pokrytia, pričom rôznym pokrytiam množiny $\overline{M}_f$ odpovedajú rôzne iredundantné d. n. f. funkcie f.

Teraz na základe uvedených vlastností určíme dolný odhad pre počet iredundantných DNF funkcie f.

Veta 4.6. Označme $\tau(f)$ počet iredundantných disjunktívnych normálnych foriem funkcie f. Nech $\epsilon_1(n) = \frac{(\lg n+1)(\lg \log_b n+2)}{2^n}$. Potom

$$\tau(f) \geq 2^{2^n \cdot (1-\epsilon_1(n))}.$$

Dôkaz. Dolný odhad pre počet iredundantných disjunktívnych normálnych foriem funkcie f získame vďaka predchádzajúcemu algoritmu.

Pretože body množiny $\overline{M}_f$ nepatria jadrovým intervalom, cez každý bod z $\overline{M}_f$ prechádzajú aspoň dva maximálne intervaly. Ďalej v dôsledku toho, že vzdialenosť medzi bodmi $\overline{M}_f$ je aspoň $2\mu+1$, každý z týchto intervalov (jeho rozmer nie je väčší ako μ) pokrýva len jeden bod z $\overline{M}_f$. Z toho vyplýva, že pokrytie vrcholov tejto množiny môžeme uskutočniť aspoň $2^{\overline{M}_f}$ spôsobmi.

Z toho už ľahko dostávame, že

$$\begin{aligned}\tau(f) &\geq 2^{\frac{|M_f|}{(2n)^{\lg\log_b n+2}}} \geq 2^{\frac{2^n(1-o(1))}{(2n)^{\lg\log_b n+2}}}\\ &\geq 2^{2^{n-(\lg+1)(\lg\log_b n+2)+\lg(1-o(1))}}\\ &\geq 2^{2^n(1-\epsilon_1(n))}.\end{aligned}$$

□

4.2 Horný odhad počtu iredundantných DNF

Na získanie horného odhadu budeme uvažovať funkcie vyhovujúce nasledujúcim podmienkam:

1. Rozmer maximálnych intervalov nie je väčší ako $\mu = \lfloor \log n - \log\log \frac{1}{p_n} \rfloor$.
2. Počet maximálnych intervalov rozmeru k vyhovuje nerovnosti
$$I_k(f) < nI_k^-(n),\ k = 0, 1, \ldots, \mu.$$
3. $I_{k_0}^+(f) = o(2^n)$, $k_0 = 0, 1, \ldots, \mu$.

Poznamenajme, že iredundatná DNF funkcie f je jednoznačne definovaná, ak

(a) je ukázaná množina intervalov rozmeru väčšieho ako k_0, ktoré vystupujú v inej (počet takých množín nie je väčší ako $2^{I_{k_0}^+(f)}$).

(b) pre každý vrchol z N_f je ukázaný jeden z intervalov rozmeru nie väčšieho ako k_0, ktorými je vrchol pokrývaný (alebo je ukázané, že vrchol nie je pokrytý takými intervalmi). Pretože cez každý vrchol prechádza nie viac ako $\sum_{k\leq k_0} c_n^k \leq n^{k_0+1}$ maximálnych intervalov, rozmer ktorých neprevyšuje k_0, môže to byť uskutočnené nie viac ako $(n^{k_0+1}+1)^{2^n}$ spôsobmi.

Označme rovnako ako v predchádzajúcej podkapitole $\tau(f)$ počet iredundantných DNF funkcie f. Z predchádzajúcich dvoch odstavcov dostávame, že

$$\tau(f) \leq 2^{I^+_{k_0}(f)} \cdot (n^{k_0+1}+1)^{2^n}.$$

Veta 4.7. Pre horné ohraničenie počtu iredundantných disjunktívnych formúl funkcie f platí nasledujúci vzťah:

$$\tau(f) \leq 2^{2^n(c\lg n+\epsilon_2(n))},$$

kde $\epsilon_2(n) = \frac{\lg 2^n+1}{2^n} + o(1)$ a $c = k_0+1$.

Dôkaz. Dôkaz tohto tvrdenia dostávame pomocou tvrdenia uvedeného pred vetou a jeho ekvivalentných úprav.

$$\begin{aligned}
\tau(f) &\leq 2^{o(2^n)} \cdot (2^{(k_0+1)\lg n}+1)^{2^n} \\
&\leq 2^{o(2^n)} \cdot \sum_{k=0}^{2^n} \binom{2^n}{k} (2^{(k_0+1)\lg n})^k \\
&\leq 2^{o(2^n)} \cdot (2^n+1) \binom{2^n}{2^{n-1}} (2^{(k_0+1)\lg n})^{2^n} \\
&\leq 2^{o(2^n)} \cdot 2^{\lg(2^n+1)} \cdot 2^{2^n} \cdot 2^{2^n(k_0+1)\lg n} \\
&\leq 2^{2^n\cdot(c\lg n+\epsilon_2(n))},
\end{aligned}$$

čo sme chceli dokázať. □

Nasledujúca veta je hlavným výsledkom tejto sekcie, pričom sumarizuje výsledky predchádzajúcich dvoch viet.

Veta 4.8. Nech $\epsilon_1(n) = \frac{(\lg n+1)(\lg\log_b n+2)}{2^n}$, $\epsilon_2(n) = \frac{\lg 2^n+1}{2^n} + o(1)$ a $c = k_0 + 1$. Potom nasledujúca nerovnosť platí s pravdepodobnosťou idúcou k 1 pre $n \to \infty$:

$$2^{2^n\cdot(1-\epsilon_1(n))} \leq \tau(f) \leq 2^{2^n(c\lg n+\epsilon_2(n))}.$$

Dôkaz. Dôkaz tejto vety triviálne vyplýva z vety 4.6 a 4.7. □

Záver

V práci sme sa zaoberali náhodne indukovanými podgrafmi hyperkocky. V našom modeli nadobúda náhodná booleovská funkcia hodnotu 1 s pravdepodobnosťou p a hodnotu 0 s pravdepodobnosťou $1 - p$ nezávisle od ostatných vrcholov. Skúmali sme štruktúry súvisiace s problémom vrcholového pokrytia takéhoto náhodného grafu a venovali sme sa počtu iredundantných disjunktných foriem náhodnej booleovskej funkcie.

Zistili sme, že pre počet iredundantných disjunktných normálnych foriem $\tau(f)$ náhodnej booleovksej funkcie f platia nasledujúce vzťahy:

$$\tau(f) \geq 2^{2^n \cdot (1-\epsilon_1(n))},$$

kde $\epsilon_1(n) = \frac{(\lg n+1)(\lg \log_b n+2)}{2^n}$, a taktiež

$$\tau(f) \leq 2^{2^n (c \lg n+\epsilon_2(n))},$$

kde $\epsilon_2(n) = \frac{\lg 2^n+1}{2^n} + o(1)$ a $c = k_0 + 1$. Vo štvrtej kapitole uvádzame výsledky, ako aj spôsob ich získania, pre ohraničenie počtu iredundantných DNF pre náhodné booleovské funkcie (vety 4.6, 4.7, 4.8).

Zaoberali sme sa aj problematikou určovania prahových funkcií. V druhej kapitole uvádzame vetu 2.4, ktorá určuje tvar prahovej funkcie pre vlasnosť obsahovania maximálnej k-rozmernej podkocky. V tretej kapitole sme za účelom získania prahovej funkcie pre vlasnosť obsahovania jadrovej k-rozmernej podkocky skúmali strednú hodnotu a disperziu náhodnej premennej C_k definovanej ako počet k-rozmerných jadrových podkociek v náhodnom grafe G. Toto úsilie však neviedlo k získaniu želanej prahovej funkcie, nakoľko sme s pomocou nám známych metód nedokázali ohraničiť hodnotu disperzie dostatočne tesne pre použitie metódy popísanej v druhej kapitole.

Z hľadiska ďalšej práce v nastolenej tematike považujeme za zaujímavé skúmať niektoré parametre náhodných booleovských funkcií pre pravdepodobnosť $p(n)$ idúcu k 0 pre n idúce do nekonečna.

Literatúra

[1] Diestel R.: Graph Theory, 3rd Ed., Springer, New York 2005.

[2] Harman R.: Pravdepodobnosť a štatistika, Poznámky k prednáškam, FMFI UK 2007.

[3] Trojaková L.: Určovanie prahových funkcií podgrafov niektorých tried grafov, Bakalárska práca, FMFI UK 2009.

[4] Toman E.: Efektívnosť niektorých lokálnych algoritmov zjednodušenia náhodných booleovských funkcií, Acta Mathematica Universitatis Comenianae, XLVIII – XLIX – 1986 (v ruštine).

[5] Toman E., Baník Ľ., Stanek M.: On structural properties of random subgraphs of n-cube, Journal of Applied Mathematics, Statistics and Infromatics, 3 (2007), No. 1.

[6] Toman E., Stanek M.: On the number of regular vertices and kernel subgraphs in random graphs, Journal of Applied Mathematics, Statistics and Infromatics, 2 (2006), No. 1.

[7] Škoviera M.: On the minimization of random boolean functions, Part 1, Computers and Artificial Inteligence, 5 (1986), No. 4.

[8] Škoviera M.: On the minimization of random boolean functions, Part 2, Computers and Artificial Inteligence, 5 (1986), No. 6.

[9] Bose, R. C., Ray-Chaudhuri, D. C.: On a class of error correcting binary group codes, Information and Control 3, 1960.

[10] Kliman J.: Odhady veľkostí pokrytí náhodne indukovaných podgrafov n-rozmernej hyperkocky, Diplomová práca, FMFI UK 2009.

[11] Švantner J.: Vlastnosti náhodne indukovaných podgrafov k-partitného grafu, Diplomová práca, FMFI UK 2006.

[12] Zvolenská Z.: Vlastnosti náhodne indukovaných podgrafov bipartitného grafu, Diplomová práca, FMFI UK 2004.