Univerzita Komenského v Bratislave
Fakulta Matematiky, Fyziky a Informatiky

Niektoré metrické vlastnosti čiastočných náhodných booleovských funkcií

Diplomová práca

2013 **Bc. Jakub Husár**

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

Niektoré metrické vlastnosti čiastočných náhodných booleovských funkcií

(Diplomová práca)

Študijný program:	Informatika
Študijný odbor:	2508 Informatika
Školiace pracovisko:	Katedra informatiky FMFI
Školiteľ:	doc. RNDr. Eduard Toman, CSc.

Bratislava, 2013 **Bc. Jakub Husár**

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Jakub Husár
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: Niektoré metrické vlastnosti čiastočných náhodných booleovských funkcií

Cieľ: Získať dolné a horné asymptotické odhady niektorých parametrov čiastočných náhodných booleovských funkcií aplikáciou kombinatoricko-pravdepodobnostných metód.

Vedúci: doc. RNDr. Eduard Toman, CSc.
Katedra: FMFI.KI - Katedra informatiky
Vedúci katedry: doc. RNDr. Daniel Olejár, PhD.

Dátum zadania: 11.09.2012

Dátum schválenia: 17.09.2012

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

.....
študent

.....
vedúci práce

Čestne prehlasujem, že túto diplomovú prácu som vypracoval samostatne s použitím uvedenej literatúry a zdrojov.

. . . . . .

Poďakovanie

Ďakujem môjmu vedúcemu diplomovej prácei doc. RNDr. Eduardovi Tomanovi, CSc. za pomoc pri výbere témy a cenné rady a pripomienky počas vypracovávania tejto práce. Ďalej ďakujem svojej snúbenici Lucii Rehákovej za podporu, ktorou ma zahŕňala a taktiež svojej rodine za trpezlivosť a porozumenie počas písania práce.

Abstrakt

Autor: Bc. Jakub Husár
Názov práce: Niektoré metrické vlastnosti čiastočných náhodných booleovských funkcií
Typ práce: Diplomová práca
Škola: Univerzita Komenského v Bratislave
Fakulta: Fakulta matematiky, fyziky a informatiky
Katedra: Katedra informatiky
Školiteľ: doc. RNDr. Eduard Toman, CSc.
Bratislava, 2013

V našej práci sa zaoberáme štúdiom niektorých merických vlastností čiastočných náhodných booleovských funkcií. Prezentujeme v nej a dokazujeme tvrdenia o asymptotických odhadoch počtu jadrových hrán a regulárnych vrcholov obsiahnutých v grafoch, ktoré sú geometrickou reprezentáciou týchto funkcií. Na základe týchto výpočtov odhadujeme zložitosť Quinovej DNF a DNF typu $\sum$ T. Tieto výsledky porovnávame so známymi zisteniami o zložitosti skrátenej DNF.

Kľúčové slová: čiastočná náhodná boolovská funkcia, jadrové hrany, regulárne vrcholy, Quinova DNF, DNF typu $\sum$ T

Abstract

Author: Bc. Jakub Husár
Thesis title: Some metric properties of partially defined random boolean functions
Thesis type: Master thesis
School: Comenius University in Bratislava
Faculty: Faculty of Mathematics, Physics and Informatics
Department: Department of Computer Science
Advisor: doc. RNDr. Eduard Toman, CSc.
Bratislava, 2013

In this work we deal with the study of some properties of partially defined random boolean functions. We are presenting and proving the theorems that asymptotically estimate the number of kernel subcubes and regular vertices contained in hypercubes equivalent to this functions. Based on these results, we estimate the complexity of Quine DNF and $\sum$ T DNF. Finally we are comparing these results with the known findings about the complexity of shortened DNF

Keywords: partially defined random boolean function, kernel subcubes, regular vertices, Quine DNF, $\sum$ T DNF

Obsah

Zoznam obrázkov

Zoznam tabuliek

Kapitola 1

Úvod

Na začiatku tejto kapitoly uvádzame fakty týkajúce sa problematiky booleovských funkcií s hlavným zameraním na čiastočné náhodné booleovské funkcie, ako aj ich možné využitie v praxi. Následne na to stanovujeme cieľ výskumu a popisujeme základné členenie našej práce.

1.1 O problematike

Keďže začiatok výskumu problematiky booleovských funkcií sa datuje ešte do začiatku 2. polovice minulého storočia, táto oblasť je pomerne dobre zmapovaná a existuje v rámci nej množstvo štúdií. Nakoľko sa v našej diplomovej práci venujeme náhodným funkciám a im zodpovedajúcim náhodným grafom, sú pre nás zaujímavé tie práce, ktoré pojednávajú o nejakej vlastnosti alebo charakteristike daných funkcií.

Medzi prvé práce tohto druhu patrí jednak práca Webera a Yablonskeho, ktorých predmetom záujmu bolo štúdium minimalizácie DNF a taktiež práca Glagoleva a Saphozenka, ktorí skúmali vlastnosti náhodných boolovských funkcií s explicitne danou pravdepodobnosťou.
Problematikou jadrových hrán a regulárnych vrcholov, o ktorých pojednávame

v našej práci sa zaoberali Toman a Stanek v práci [9], odkiaľ pochádzajú aj niektoré pre našu zaujímavé výsledky ohľadom štúdia vlastností úplných náhodných booleovských funkcií.

Menej prebádanú oblasť tvorí oblasť venujúca sa špeciálnemu typu booleovských funkcií, ktoré nemusia byť definované na celom vstupe, jedná sa o tzv. čiastočné booleovské funkcie.
Výskumu v tejto oblasti sa venovala aj L. Haviarová v práci [5], ktorá preskúmavala počty hrán a maximálnych hrán v náhodnom grafe reprezentujúcom čiastočnú náhodnú booleovskú funkciu a z toho vyplývajúce zložitosti úplnej a skrátenej DNF. Na výsledky tejto práce nadväzujeme v našom výskume.

1.2 Využitie

Okrem samotného matematického hľadiska majú výsledky zo skúmanej problematiky uplatnenie aj v rôznych oblastiach informatiky. Vieme ich využiť či už pri návrhu a vytváraní logických obvodov, kde každý fyzický elektrický obvod zodpovedá DNF funkcii a teda jej minimalizovanie má vplyv na zložitosť týchto obvodov. Taktiež v kryptológii, kde koncept čiastočných náhodných booleovských funkcií môže byť využitý na vytváranie kryptograficky silných booleovských funkcií, ktoré majú využitie napríklad pri vytváraní efektívnych šifrovacích algoritmov alebo hašovacích funkcií. Rovnako tak vieme výsledky zužitkovať aj pri snahe o rozpoznávanie obrazcov, vytváraní testov a taktiež sa dajú poznatky využiť aj pri štúdiu problematiky oddeliteľnosti množín.
O problematike oddeliteľnosti množín a teórii testov je bližšie uvádzané v nasledujúcej kapitole.

1.3 Cieľ

Na začiatku skúmania danej problematiky bolo potrebné stanoviť si niekoľko cieľov, ktoré by udávali smerovanie celej našej práce. Jedná sa konkrétne o tieto ciele:

1. asymptoticky odhadnúť a ohraničiť počet jadrových hrán čiastočných booleovských funkcií
2. asymptoticky odhadnúť a určiť dolnú a hornú hranicu pre počet regulárnych vrcholov čiastočných booleovských funkcií
3. vyjadriť zložitosti Quinovej DNF a DNF typu $\sum T$ a porovnať ich so zložitosťou skrátenej DNF z práce [5]

1.4 Členenie práce

Túto prácu sme rozčlenili do jednotlivých kapitol.
Definície základných pojmov a označení, ktoré v práci používame: čiastočná booleovská funkcia, rôzne druhy DNF, geometrický prístup k problematike, všeobecný pravdepodobnostný model a použité metódy tvoria druhú kapitolu. Okrem toho na tomto mieste, uvádzame znenie problémov, na ktoré vieme aplikovať náš výskum.
V tretej kapitole predstavujeme vlastné výsledky. Určujeme strednú hodnotu a ohraničenia počtu jadrových hrán, ako aj strednú hodnotu a hornú hranicu počtu regulárnych vrcholov. Všetky tvrdenia, ktoré uvádzame zároveň náležitým spôsobom dokazujeme. Na základe týchto tvrdení určujeme zložitosti Quinovej DNF a DNF typu $\sum T$.
V závere práce sumarizujeme a hodnotíme všetky dosiahnuté výsledky a zamýšľame sa nad možnosťami ďalšieho výskumu.

Kapitola 2

Základné definície a pojmy

V tejto kapitole uvedieme základné pojmy, definície a označenia, ktoré budeme v ďalšom texte používať.
Najskôr zadefinujeme, čo rozumieme pod pojmom čiastočná booleovská funkcia (ČBF), ďalej uvedieme definíciu disjunktívnej normálnej formy a to, ako pomocou nej vieme realizovať čiastočnú booleovskú funkciu. Aby sme mohli pracovať s ČBF, zadefinujeme pojmy týkajúce sa pravdepodobnosti. Taktiež uvedieme metódy, ktoré budeme využívať na vyčíslovanie jednotlivých vlastností a napokon udáme označenia, s ktorými budeme v ďalšom texte pracovať.

2.1 Čiastočná booleovská funkcia

Pod pojmom čiastočná booleovská funkcia rozumieme také zobrazenie, ktoré pre množinu B_{def}, ktorá je podmnožinou množiny $\{0,1\}^n$, priradí hodnotu z množiny $B = \{0,1\}$, pričom pre každú n-ticu z $B^n \setminus B_{DEF}$ platí, že funkcia na tomto vstupe nie je definovaná.
Táto funkcia preto na definovaných vstupoch nadobúda tvar

$$B_{def} \to B$$

Inými slovami, je to taká funkcia, ktorá vybraným n-ticiam premenných $x_1, x_2, ..., x_n$ priradí hodnotu z B a pre zvyšné n-tice nie je táto funkcia definovaná.

Nakoľko máme 2^n rôznych n-tíc vstupov a tie vieme rozdeliť do 3 rôznych množín podľa toho, či funkcia na danom vstupe nadobúda hodnotu 0, 1, alebo nie je definovaná, dostávame, že existuje práve 3^{2^n} rôznych čiastočných booleovských funkcií.

Príklad 2.1.1 Príklad čiastočnej booleovskej funkcie

x_1	x_2	x_3	x_4	$f(x_1, x_2, x_3, x_4)$
0	0	0	0	0
0	0	1	0	1
0	1	0	0	1
0	1	0	1	0
0	1	1	1	1
1	0	0	1	1
1	0	1	0	1
1	1	0	1	1
1	1	1	0	0
1	1	1	1	1

Tabuľka 2.1: Čiastočná booleovská funkcia daná pravdivostnou tabuľkou

Funkcia f je funkciou o štyroch premenných, kde definičný obor tvorí množina $\{0000, 0010, 0100, 0101, 0111, 1001, 1010, 1101, 1110, 1111\}$ a funkcia nie je definovaná na zvyšných vstupoch, teda vstupoch z množiny $\{0001, 0011, 0110, 1000, 1011, 1100\}$. Obor hodnôt tejto funkcie je daný množinou $B = \{0, 1\}$.

Označenie 2.1.1 *Nech f je n-árna čiastočná booleovská funkcia. Množinou N_f označíme takú množinu n-tíc $(a_1, a_2, ..., a_n) \in B^n$, pre ktoré platí, že*

$$(a_1, a_2, ..., a_n) \in N_f$$

vtedy a práve vtedy, keď

$$f(a_1, a_2, ..., a_n) = 1$$

Obdobne budeme označovať množinou $N_{\bar{f}}$ tie n-tice, pre ktoré $f(a_1, a_2, ..., a_n) = 0$ a množina $N_{\tilde{f}}$ bude reprezentovať tie, pre ktoré nie je funkcia f definovaná.

2.1.1 Realizácia disjunktívnou normálnou formou

Najprv zavedieme niekol'ko užitočných pojmov, ktoré budeme pri popise a disjunktívnych normálnych foriem používat'.

Označenie 2.1.2

$$x^\sigma = x\sigma \vee \bar{x}\bar{\sigma},\ kde\ \sigma \in \{0, 1\}$$

Potom platí

$$x^\sigma = \begin{cases} \bar{x} & ak\ \sigma = 0 \\ x & ak\ \sigma = 1 \end{cases}$$

čo sa dá vyjadriť ako

$$x^\sigma = \begin{cases} 1 & ak\ x = \sigma \\ 0 & ak\ x \neq \sigma \end{cases}$$

Výraz $x_i^{\sigma_i}$ ktorý predstavuje premennú x_i alebo jej negáciu $\bar{x}_i$ budeme nazývat' literálom.

Definícia 2.1.1 (Elementárna konjunkcia) *Nech $x_1, \ldots, x_n$ sú premenné nejakej n-árnej booleovskej funkcie a nech $x_1^{\sigma_1}, \ldots, x_r^{\sigma_r}$, kde $0 < r < n$, sú*

l'ubovol'né literály týchto premenných. Potom výraz $K = x_1^{\sigma_1} \wedge \ldots \wedge x_r^{\sigma_r}$ *budeme nazývat' elementárnou konjunkciou. Hodnotu r nazývame rádom konjunkcie K a označovať ju budeme* $r(K)$*, pričom konjunckiu K s* $r(K) = 0$ *nazývame prázdnou konjunkciou a jej hodnota je konštantne rovná 1.*

Definícia 2.1.2 (Disjunktívna normálna forma) *Nech* $K_1, \ldots, K_m$ *sú navzájom rôzne elementárne konjunkcie. Potom výraz* $D = K_1 \vee \ldots \vee K_m$ *nazývame disjunktívnou normálnou formou (DNF). Hodnota m sa nazýva dĺžkou DNF, pričom ak* $m = 0$*, tak hovoríme, že DNF D je prázdna a má hodnotu 0.*

Definícia 2.1.3 (DNF realizujúca funkciu f) *Hovoríme, že DNF D realizuje čiastočnú booleovskú funkciu f práve vtedy, keď platí, že pre všetky vstupy* $(a_1, \ldots, a_n)$*,* $a_i \in \{0, 1\}$ *na ktorých je funkcia f definovaná, je hodnota* $D(a_1, \ldots, a_n)$ *totožná s hodnotou* $f(a_1, \ldots, a_n)$*. (Mimo oblasti definície funkcie f, môže DNF D nadobúdať ľubovoľné hodnoty).*
Inými slovami, ak N_D *je množina všetkych n-tíc, pre ktoré je D splnená, potom DNF D realizuje funkciu f práve vtedy keď platí*

$$(N_f \subseteq N_D) \wedge (N_{\bar{f}} \cap N_D = \emptyset)$$

Každú čiastočnú booleovskú funkciu vieme realizovať pomocou disjunktívnej normálnej formy (DNF). Jednej čiastočne booleovskej funkcii môže ale zospovedať niekoľko rôznych DNF, viď nasledujúci príklad.

Príklad 2.1.2 Uvažujme terárnu čiastočnú booleovskú funkciu $f(x_1, x_2, x_3)$ zadanú nasledujúcou pravdivostnou tabuľkou

Funkcia f je realizovaná disjunktívnou normálnou formou D_1

$$D_1 = \bar{x_1} x_2 x_3 \vee x_1 \bar{x_3} \vee x_2 \bar{x_3}$$

x_1	x_2	x_3	$f(x_1, x_2, x_3)$
0	0	1	0
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	0
1	1	0	1

Tabuľka 2.2: Funkcia f ako príklad nejednoznačnosti DNF

avšak, taktiež môže byť realizovaná aj pomocou DNF D_2

$$D_2 = x_2x_3 \vee \bar{x_3}$$

Práve pre túto nejednoznačnosť potebujeme uviesť, ktoré DNF realizujúce funkciu f nás budú v nasledujúcom texte zaujímať.

Existuje viacero kritérií, pomocou ktorých môžeme posudzovať tieto DNF. Aby sme ich určili zadefinujeme si nasledujúce pojmy.

Definícia 2.1.4 (Implikant) *Elementárna konjunkcia K sa nazýva implikantom funkcie f, ak existuje n-tica $(a_1, \ldots, a_n) \in N_f$ pre ktorú $K(a_1, \ldots, a_n) = 1$ a zároveň platí, že pre každú n-ticu $(b_1, \ldots, b_n) \in N_{\bar{f}}$, $K(b_1, \ldots, b_n) = 0$.*

Definícia 2.1.5 (Prostý implikant) *Prostým implikantom funkcie f nazývame implikant K taký, že ľubovoľná konjunkcia ktorú získame z K vynechaním niektorého činiteľa ${x_i}^{\sigma_i}$, nie je už implikantom funkcie f.*

Teraz si uvedieme niekoľko typov DNF, ktoré budeme v nasledujúcom texte využívať.

Definícia 2.1.6 (Skrátená DNF) *Skrátenou DNF funkcie f budeme nazývať takú DNF D, ktorá je disjunkciou všetkých prostých implikantov funkcie f a označovať ju budeme $D_s(f)$.*

Poznámka 2.1.1 Ku každej čiastočnej booleovskej funkcii f existuje práve 1 skrátená DNF.

Definícia 2.1.7 (Iredundantná DNF) *Nech D je DNF realizujúca funkciu f a D sa skladá iba zo samých prostých implikantov funkcie f. Potom, ak z D vynechaním ľubovolného prostého implikantu K_i, D' už nerealizuje funkciu f, nazývame túto DNF iredundantnou.*

Aby sme mohli zadefinovať ďalšie typy DNF, je potrebné uviesť, čo musia spĺňať parametre charakterizujúce zložitosti týchto DNF, na základe ktorých budeme tieto DNF porovnávať. Preto teraz definujeme pojem index jednoduchosti L(D), ktorý musí spĺňať niekoľko axióm.

Definícia 2.1.8 (Index jednoduchosti)
Indexom jednoduchosti DNF D nazývame taký parameter L(D), ktorý spĺňa všetky nasledujúce axiómy.

1. *Axióma nezápornosti. Pre ľubovoľnú DNF $L(D) \geq 0$.*
2. *Axióma monotónnosti (vzhľadom na násobenie). Nech $D = D' \vee x_i^{\sigma_i} K_i'$. Potom $L(D) \geq L(D' \vee K_i')$.*
3. *Axióma vypuklosti (vzhľadom na sumáciu). Nech $D = D_1 \vee D_2$. Ak $D_1 \wedge D_2 = 0$, tak platí $L(D) \geq L(D_1) + L(D_2)$.*
4. *Axióma invariantnosti (vzhľadom na izomorfizmus). Nech DNF D' bola získaná z DNF D premenovaním premenných (bez toho aby boli stotožnené). Potom $L(D) = L(D')$.*

Teraz uvedieme 2 typy DNF, ktoré sa určujú na základe ich zložitosti a ktorých zložitosť spĺňa pomienky indexu jednoduchosti.

Definícia 2.1.9 (Minimálna DNF) *Iredundantná DNF s minimálnou zložitosťou (minimálnym počtom činiteľov) sa nazýva minimálna DNF. Tento počet činiteľov budeme v ďalšom texte označovať ako L(D).*

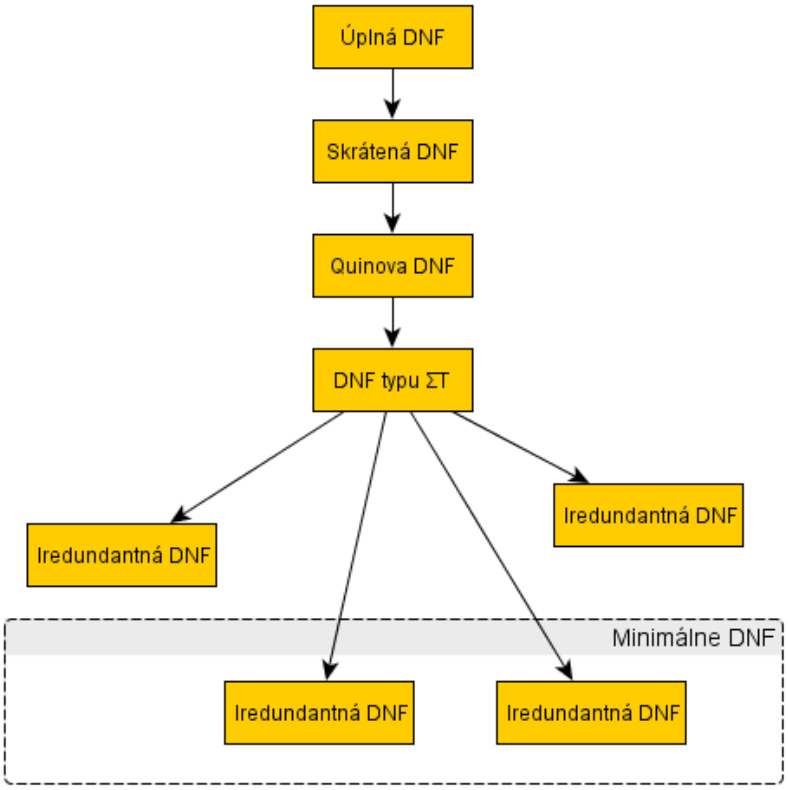
Definícia 2.1.10 (Najkratšia DNF) *Iredundantná DNF s minimálnou dĺžkou (minimálnym počtom implikantov) sa nazýva najkratšia DNF a túto dĺžku budeme označovať ako l(D).*

V pôvodnom uvažovaní procesu minimalizácie sa iredundantné DNF odvodzovali priamo zo skrátenej DNF. Nakoľko ale táto časť tohto procesu, kde sa proces vetví a vytvárajú sa iredundantné DNF je najnáročnejšia, pokúšame sa ho zjednodušiť tým, že sa vynasnažíme vopred vylúčiť časť členov skrátenej DNF, ktoré sa nezúčastňujú pri zostrojovaní iredundantných DNF, a tým skrátiť celkové prezeranie. Realizujeme to pritom tak, aby zvyšná časť členov umožnila zostrojiť aspoň jednu minimálnu DNF. Dostávame tak nový proces minimalizácie, ktorý je uvedený na obrázku 2.1.
Najvhodnejšie je, aby tento krok bol realizovaný jednoznačne.

Preto zavedieme Quinovu DNF a DNF typu $\sum T$, ktoré nám tieto kroky budú realizovať.

Definícia 2.1.11 (Quinova DNF) *Majme množinu tých prostých implikantov, ktoré sú obsiahnuté v každej iredundantnej DNF. Tieto implikanty nazvyme jadrové implikanty. Potom DNF D, ktorú dostaneme zo skrátenej DNF $D_s(f)$ vynechaním všetkých tých prostých implikantov, ktorých množina prípustných hodnôt je pokrytá jadrovými implikantmii, nazývame Quinova DNF funkcie f a označovať ju budeme N_Q.*

Definícia 2.1.12 (DNF typu ΣT) *DNF D, ktorá obsahuje všetky tie prosté implikanty, ktoré sú obsiahnuté v krajnom prípade v aspoň jednej iredundantnej DNF funkcie f nazývame DNF typu ΣT tejto funkcie a označovať ju budeme $N_{\Sigma T}$*



Obr. 2.1: Proces minimalizácie DNF

2.1.2 Geometrická reprezentácia

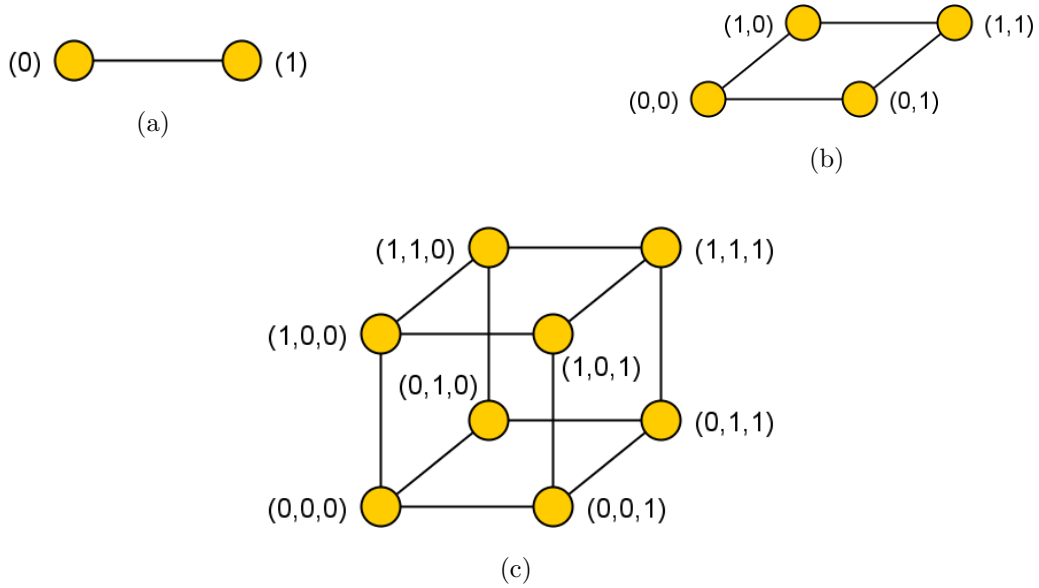
Čiastočnú booleovskú funkciu f vieme reprezentovať aj pomocou geometrickej formy, a totiž grafu.

Na množinu všetkých možných binárnych n-tíc B^n môžme tiež pozerať ako na množinu všetkých vrcholov n-rozmernej jednotkovej kocky(tiež nazývanej ako *hyperkocka*), kde n-tice $(a_1, a_2, ..., a_n)$ reprezentujú vrcholy tejto kocky a 2 vrcholy sú spojené hranou práve vtedy, keď sa líšia vo svojom označení v práve 1 súradnici.

Aby sme sa vyhli komplikáciam s izomorfizmom, každému vrcholu takejto

kocky priradíme jedinečné označenie v podobe príslušného binárneho vektora.

Príklad 2.1.3 Jednotkové kocky rozmeru 1, 2 a 3



Obr. 2.2: Príklad jednotkových kociek, kde (a) reprezentuje unárnu, (b) binárnu a (c) ternárnu booleovskú funkciu

Definícia 2.1.13 (Hrana) *Nech $\sigma_{i_1}, \sigma_{i_2}, ..., \sigma_{i_k}$ je nejaká pevne zvolená k-tica čísel z množiny $\{0, 1\}$, pre ktoré platí, že $1 \leq i_1 < i_2 < ... < i_k \leq n$. Množinu všetkých vrcholov $(a_1, a_2, ..., a_n)$ kocky B^n takých, že $a_{i_1} = \sigma_{i_1}, a_{i_2} = \sigma_{i_2}, ..., a_{i_k} = \sigma_{i_k}$ nazývame (n-k)-rozmerná hrana.*

Definícia 2.1.14 (Interval) *Nech $K(x_1, x_2, ..., x_n) = x_{i_1}^{\sigma_{i_1}} \wedge x_{i_2}^{\sigma_{i_2}} \wedge ... \wedge x_{i_k}^{\sigma_{i_k}}$ je elementárna konjunkcia dĺžky k. Množinu vrcholov grafu B^n takých, ktoré majú na pozíciach $j \in \{1, 2, \ldots, k\}$ hodnoty 1 v prípade že σ_{i_j} je rovné 1 a hodnoty 0 keď $\sigma_{i_j} = 0$, nazývame interval k-tého rádu a označovať ho budeme N_K.*

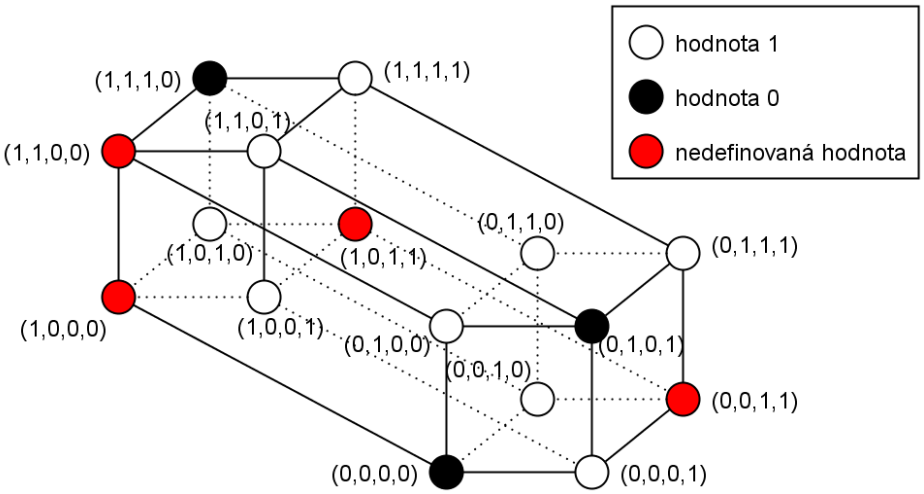
Interval k-tého rádu N_K zodpovedá $(n-k)$-rozmernej hrane kocky B^n.

Definícia 2.1.15 (Maximálny interval) *Interval N_K odpovedajúci konjunkcii K nazývame maximálnym, ak neexistuje žiaden interval N'_K, taký, že*

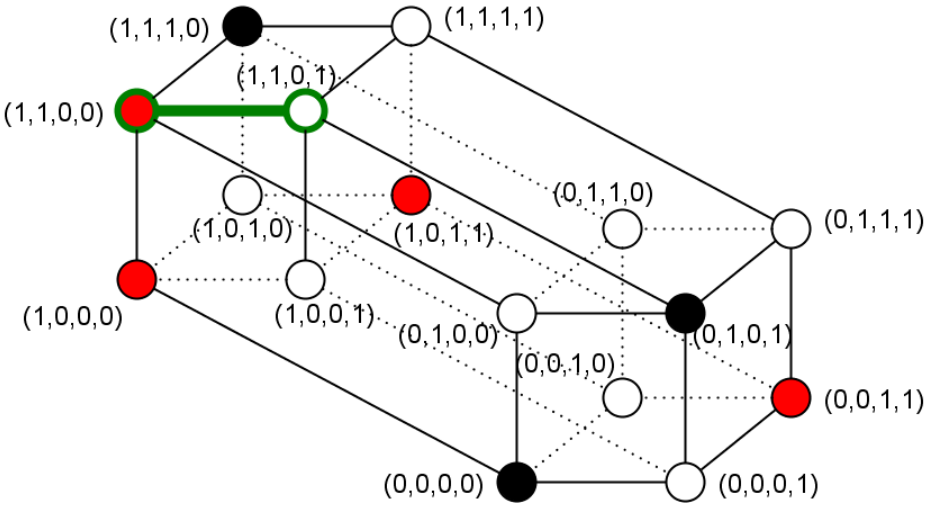
1. $N_K \subseteq N'_K \subseteq (N_f \cup N_{\tilde{f}})$
2. *Rád intervalu N'_K je menší ako rád intervalu N_K*

Poznámka 2.1.2 Interval N_K budeme v ďalšom texte v súvislosti s kockou B^n označovať aj pojmom *hrana* a maximálny interval tiež ako *maximálna hrana*.

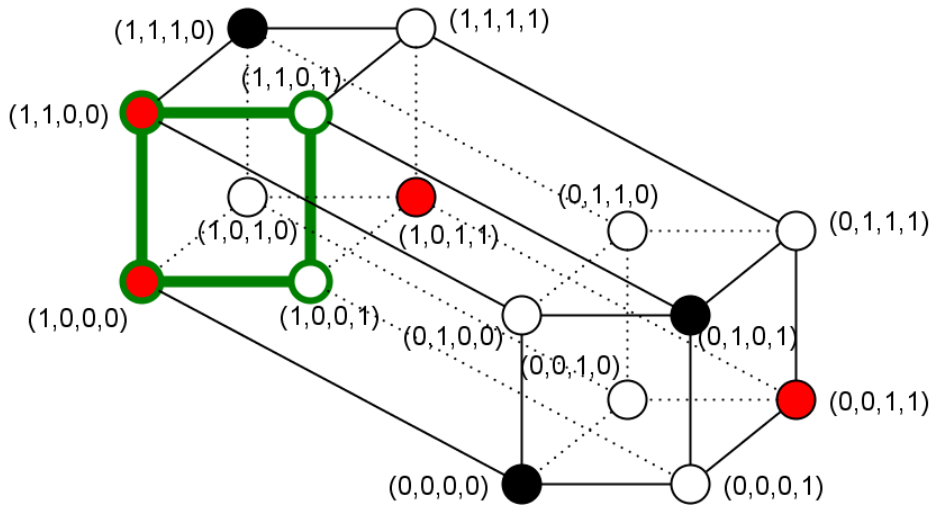
Príklad 2.1.4 Na nasledujúcich obrázkoch uvádzame príklad hyperkocky reprezentujúcej čiastočnú booleovskú funkciu, ako aj príklad intervalu a maximálneho intervalu v tejto hyperkocke. Čiastočná booleovská funkcia, ktorú realizujeme je prebratá z príkladu 2.1.1.
Počnúc týmto miestom ďalej budeme vrcholy s hodnotou 1 značiť bielou farbou, vrcholy s hodnotou 0 čiernou farbou a nedefinované vrcholy budú označované červenou farbou. Vrcholy v intervaloch a ich prepojenia budeme označovať zelenou farbou.



Obr. 2.3: 4-rozmerná jednotková kocka reprezentujúca čiastočnú booleovskú funkciu z príkladu 2.1.1



Obr. 2.4: Interval N_K 3-tieho rádu zodpovedajúci konjunkcii $K = x_1^1 \wedge x_2^1 \wedge x_3^0$, kde $N_K = \{(1,1,0,0),(1,1,0,1)\}$. Tento interval však nie je maximálny, pretože existuje interval N'_K, ktorý obsahuje všetky vrcholy z N_K a ktorého rád je zároveň menší ako rád intervalu N_K



Obr. 2.5: Maximálny interval $N_{K'}$ rádu 2, ktorý zodpovedá konjunkcii $K' = x_1^1 \wedge x_3^0$ a ktorému prislúchajú vrcholy {(1,0,0,0), (1,0,0,1), (1,1,0,0), (1,1,0,1)}

2.2 Pravdepodobnosť

2.2.1 Definícia pravdepodobnostného priestoru

Uvažujme o nejakom náhodnom jave. Uskutočnenie tohto náhodného javu budeme nazývať náhodná udalosť a budeme to značiť ω. Množinu všetkých elementárnych náhodných udalostí budeme nazývať priestor elementárnych udalostí a značiť ju budeme Ω.

Prvky množiny Ω budeme nazývať elementárne výsledky.

Definícia 2.2.1 (Pole udalostí) *Neprázdny systém S podmnožín priestoru elementárnych udalostí Ω, ktorý obsahuje ako prvok Ω a je uzavretý vzhľadom na komplement a spočítateľné zjednotenie, sa nazýva pole udalostí nad výberovým priestorom Ω. Prvky takéhoto poľa S budeme nazývať udalosti.*

To znamená, že S je pole udalostí, ak spĺňa nasledujúce podmienky:

1. $\Omega \in S$
2. $\forall A; A \in S \Rightarrow A^c = \Omega - A \in S$
3. $\forall A_i; i = 1, 2, ..., n; A_i \in S \Rightarrow \bigcup_{i=1}^{n} A_i \in S$

Definícia 2.2.2 (Ω-algebra) *Nech S je pole náhodných udalostí a Ω je priestor elementárnych udalostí, z ktorého vyberáme. Usporiadanú dvojicu (Ω, S) potom nazývame Ω-algebra udalostí.*

Definícia 2.2.3 (Pravdepodobnostná miera) *Zobrazenie $P : S \to R$ nazývame pravdepodobnostná miera na Ω-algebre udalostí, práve vtedy, keď spĺňa nasledovné podmienky:*

1. $\forall A \in S$ *platí* $0 \leq P(A) \leq 1$
2. $P(\emptyset) = 0$ *a* $P(\Omega) = 1$
3. $\forall A_i; i = 1, 2, ..., n; A_i \in S$, *také že* $A_i \cap A_j = \emptyset$ *pre* $i \neq j$ *platí*

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

Definícia 2.2.4 (Pravdepodobnostný priestor) *Nech (Ω, S) je Ω-algebra udalostí a P nech je pravdepodobnostná miera na tejto Ω-algebre udalostí. Potom usporiadanú trojicu (Ω, S, P) nazývame pravdepodobnostný priestor náhodných udalostí.*

Definícia 2.2.5 (Náhodná premenná) *Nech (Ω, S, P) je nejaký pravdepodobnosntný priestor. Funkciu na reálnych číslach $X : \Omega \to R$ budeme nazývať náhodná premenná, ak platí:*

$$\forall x \in R : \{\omega; X(\omega) < x\} \in S$$

Definícia 2.2.6 (Diskrétna náhodná premenná) *Náhodná premenná X sa nazýva diskrétna, ak existuje postupnosť reálnych čísel $\{x_i\}$ a postupnosť nezáporných reálnych čísel $\{p_i\}$, takých že*

$$P(X = x_i) = p_i$$

a zároveň

$$\sum_{i \in I} p_i = 1$$

Definícia 2.2.7 (Stredná hodnota diskrétnej náhodnej premennej) *Nech X je diskrétna náhodná premenná a nech $\{x_i\}$ je postupnosť reálnych čísel, ktorej hodnoty môže s nenulovou pravdepodobnosťou nadobúdať. Potom hovoríme, že náhodná premenná X má konečnú strednú hodnotu $E(X)$ a ju vyčíslujeme ako:*

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

Definícia 2.2.8 (Vlastnosť náhodnej funkcie) *Nech A je nejaká vlastnosť a f nech je nejaká náhodná booleovská funkcia. Ak*

$$\lim_{n \to \infty} P[f \text{ má vlastnosť } A] = 1,$$

tak hovoríme, že náhodná booleovská funkcia f má vlastnosť A resp. f spĺňa vlastnosť A takmer s určitosťou.

Poznámka 2.2.1 V takomto prípade hovoríme tiež, že takmer všetky náhodné booleovské funkcie majú vlastnosť A.

2.2.2 Pravdepodobnostné metódy

V ďalšom texte budeme vo viacerých prípadoch využívať na vyčíslovanie hodnôt Markovovu pravdepodobnostnú metódu (a jej priamy dôsledok), ktorá poskytuje základné prostriedky pre konštrukciu odhadov náhodných premenných.

Definícia 2.2.9 (Markovova nerovnosť) *Nech X je náhodná premenná nadobúdajúca nezáporné hodnoty. Nech $t > 0$ je reálne číslo. Potom platí:*

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

Poznámka 2.2.2 Ako dôsledok Markovovej nerovnosti dostávame vzťah

$$P[X \geq t.E(X)] \leq \frac{1}{t}$$

2.3 Asymptotické ohraničenia a ďalšie ohraničenia

Na tomto mieste uvádzame niektoré spôsoby asymptotických ohraničení funkcií, ktoré budeme v ďalšom texte používať.

Definícia 2.3.1 (o-notácia) *Symbolom $o(a_n)$ označujeme taký výraz, ktorý keď je delený a_n ide k 0.*

Definícia 2.3.2 (O-notácia) *Symbolom $O(a_n)$ označujeme taký výraz, ktorý keď je delený a_n ostáva ohraničený nejakou pevnou konštantou $c > 0$.*

Definícia 2.3.3 (asymptotická ekvivalencia) *Hovoríme, že postupnosti (a_n) a (b_n) sú asymptoticky ekvivalentné, ak $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$. Túto skutočnosť označujeme tiež zápisom $a_n \sim b_n$.*

Symbolom $\lg n$ budeme označovať logaritmus čísla n pri základe 2. Nakoľko budeme často používať logaritmus pri základe $\frac{1}{p_1+p_3}$, pre zjednodušenie označenia položíme $b = \frac{1}{p_1+p_3}$ a namiesto $\log_{1/(p_1+p_3)} n$ budeme písať $\log_b n$.

Napokon, označením $\lfloor x \rfloor$ budeme rozumieť dolnú celú časť čísla x a $\lceil x \rceil$ jeho hornú celú časť.

2.4 Aplikácie

2.4.1 Oddeliteľnosť množín

Nech B^n je jednotková hyperkocka. Uvažujme nie všade definovanú booleovskú funkciu (čiastočnú booleovskú funkciu) danú predpisom

$$f(x_1,\ldots,x_n) = \begin{cases} 0 & \text{ak } (x_1,\ldots,x_n) \in M_0 \\ 1 & \text{ak } (x_1,\ldots,x_n) \in M_1 \\ \text{nie je definovaná} & \text{ak } (x_1,\ldots,x_n) \in B^n \backslash (M_0 \cup M_1) \end{cases}$$

kde $M_1 \cap M_0 = \emptyset$, $M_0 \subseteq B^n$, $M_1 \subseteq B^n$.

Nech $P_2(n)$ je množina všade definovaných booleovských funkcií o n-premenných.
Nech $P(f)$ je číselná funkcia na $P_2(n)$. Túto funkciu budeme nazývať funkciou zložitosti charakterizujúcou funkciu $f \in P_2(n)$.

Funkciu $f' \in P_2(n)$ nazývame prípustnou vzhľadom na f, ak pre všetky $\alpha \in M_0$: $f'(\alpha) = 0$ a pre všetky $\beta \in M_1$: $f'(\beta) = 1$.

Úloha je v tomto prípade definovaná nasledovne:
Nájsť spomedzi všetkých prípustných funkcií $f' \in P_2(n)$ takú f^*, pre ktorú funkcia zložitosti $P(f^*)$ nadobúda minimálnu hodnotu.
(úloha sa dá interpretovať vzhľadom na rôzne typy zložitostí funkcií)

2.4.2 Teória testov

Nech $f(x_1, \ldots, x_n)$ je booleovská funkcia o n premenných. $T_n(f)$ nech je množina skladajúca sa z n-rozmerných booleovských vektorov. Táto množina sa nazýva preverujúci test pre $f(x_1, \ldots, x_n)$, ak pre ľubovoľnú funkciu $\varphi(x_1, \ldots, x_n)$, ktorú dostaneme z $f(x_1, \ldots, x_n)$ zamenením niektorej premennej za konštantu 0 alebo 1 tak aby existoval vektor $u = (u_1, \ldots, u_n)$, $u \in T_n(f)$, že $f(x_1, \ldots, x_n) \neq \varphi(x_1, \ldots, x_n)$.
Počet vektorov v teste sa nazýva zložitosť testu a označuje sa $L(T_n(f))$.
Test pre funkciu $f(x_1, \ldots, x_n)$, ktorý má najmenšiu zložitosť sa nazýva minimálnym a zložitosť tohto minimálneho testu označujeme $L(f)$.
Zavedieme označenie:

$$L_K(n) = \max_{f(x_1,\ldots,x_n)\in K} L(f)$$

kde K je niektorá množina booleovských funkcií.

Alternatívna verzia:
Nech P_2^n je množina všetkých funkcií algebry logiky o n premenných, Q_2^n nech je množina čiastočne definovaných funkcií algebry logiky o n premenných, R^n nech je n-rozmerný vektorový priestor nad poľom reálnych čísel R a B^n nech je podmnožina R^n zložená z binárnych vektorov dĺžky n.
Nech $f \in Q_2^n$. Množinou E_f budeme označovať takú množinu $\alpha = (\alpha_1, \ldots, \alpha_n) \in B^n$ takých, že $f(\alpha) = 1$ a množinou N_f množinu vrcholov $\beta = (\beta_1, \ldots, \beta_n) \in B^n$ takých, že $f(\beta) = 0$. Vektor $\tau = (\tau_1, \ldots, \tau_n) \in B^n$ sa nazýva testom pre $f \in Q_2^n$, ak pre ľubovoľné $\alpha \in E_f$ a $\beta \in N_f$ platí vzťah $(\tau_1.\alpha_1, \ldots, \tau_n.\alpha_n) \neq (\tau_1.\beta_1, \ldots, \tau_n.\beta_n)$.
Test $\tau = (\tau_1, \ldots, \tau_n) \in B^n$ sa nazýva iredundantným testom pre $f \in Q_2^n$, ak pre ľubovoľné $\alpha \in B^n$, kde $\alpha < \tau$ platí, že α nie je test pre f.

Kapitola 3

Odhady niektorých metrických vlastností čiastočných booleovských funkcií

Na začiatku tejto kapitoly popíšeme pravdepodobnosntý priestor, ktorý tvoria čiastočné náhodné booleovské funkcie, s ktorými budeme v tejto práci pracovať. Následne budeme skúmať početnosť jadrových hrán a regulárnych vrcholov, obsiahnutých v grafe B^n, ktorý zodpovedá danej funkcií. Z toho napokon asymptopticky odhadneme dĺžku Quinovej DNF resp. DNF typu Σ.

3.1 Pravdepodobnostný priestor

Tak ako sme si v kapitole s definíciami zadefinovali všeobecný pravdepodobnostný priesor, tak aj čiastočné náhodné booleovské funkcie tvoria takýto diskrétny pravdepodobnostný priestor. Budeme ho označovať príslušnou trojicou $(P^n, 2^{P^n}, P)$, pričom P^n označuje množinu všetkých n-árnych čiastočných náhodných booleovských funkcií, 2^{P^n} reprezentuje všetky možné výbery, ktoré

sa môžu uskutočniť a P predstavuje pravdepodobnosť, s ktorou pracujeme pri týchto funkciách.

Náhodný výber funkcie $f \in P^n$ si môžeme predstaviť ako náhodné rozdelenie jej vstupov do troch množín: množinu tých vstupov, na ktorých má funkcia hodnotu 1; tých, na ktorých dosahuje funkcia výstup 0; a napokon tých, pre ktoré funkcia nie je definovaná. Tieto množiny si v prípade ekvivalentného grafu B^n môžme reprezentovať aj ako N_f, $N_{\bar{f}}$, $N_{\tilde{f}}$. Teraz si potrebujeme zadefinovať, s akými pravdepodobnosťami sa n-tica $\alpha \in B^n$ vyskytne v akej množine.

- p_1 - pravdepodobnosť, že pripadne množine N_f, a teda bude sa zobrazovať na hodnotu 1
- p_2 - pravdepodobnosť, že pripadne množine $N_{\bar{f}}$, a teda bude sa zobrazovať na hodnotu 0
- p_3 - pravdepodobnosť, že pripadne množine $N_{\tilde{f}}$, a teda na danom vstupe nebude funkcia f definovaná

Tieto pravdepodobnosti sú nezávislé od akýchkoľvek predchádzajúcich výberov a popritom sú zadefinované tak, že platí vzťah $p_1 + p_2 + p_3 = 1$.

Pravdepodobnosť výberu funkcie $f \in P^n$ je teda daná vzťahom

$$P[\{f\}] = {p_1}^{|N_f|} \cdot {p_2}^{|N_{\bar{f}}|} \cdot {p_3}^{|N_{\tilde{f}}|}$$

Z toho dostávame, že pre ľubovoľnú podmnožinu A čiastočných náhodných booleovských funkcií P^n platí

$$P[A] = \sum_{f \in A} P[\{f\}]$$

Nakoľko ľubovolné funkcie f_1, f_2 sú jednoznačne určené na každom zo vstupov, nemôže sa stať, že by 2 ľubovolné n-árne funkcie, ktoré by boli

rôzne, nastali v ten istý čas. Hovoríme teda, že udalosti výberu týchto funkcií sú vzájomne exkluzívne, alebo inými slovami disjunktné. Ak 2 udalosti sú disjunktné, potom pravdepodobnosť, že nastane ktorákoľvek z nich, je daná nasledujúcim pravidlom $P(A \vee B) = P(A) + P(B)$ Preto

$$P[A] = P(f_1 \vee f_2 \vee \cdots \vee f_k) = \sum_{f_i \in A} P[\{f_i\}]$$

3.2 Jadrové hrany

Snaha zjednodušiť zložitosť DNF a tým proces minimalizácie tak, aby toto zjednodušenie sa realizovalo jednoznačým spôsobom nás privádza k pojmu jadrových hrán.

3.2.1 Zadefinovanie pojmov

Jadrové hrany sú také hrany v rámci grafu G, ktoré musia byť súčasťou ľubovoľného vrcholového pokrytia grafu G, pričom toto pokrytie je realizované pomocou maximálnych hrán.

Definícia 3.2.1 *Maximálnu hranu N_K obsiahnutú v grafe G nazývame jadrovou hranou, ak existuje vrchol $v \in N_K$ taký, že tento nie je obsiahnutý v žiadnej inej maximálnej hrane, ktorá sa vyskytuje v rámci grafu G. Množinu všetkých jadrových hrán v G nazývame jadro grafu G.*

Príklad 3.2.1 Majme čiastočnú booleovskú funkciu, ktorá je reprezentovaná hyperkockou znázornenou na nižšie uvedených obrázkoch a) a b). Ako vidno z obrázku a) táto obsahuje práve 3 maximálne hrany $e1$, $e2$, $e3$ a prislúcha jej skrátená DNF $D = \bar{x_2}\bar{x_3} \wedge x_1\bar{x_2} \wedge x_1x_3$. Vrchol (0,0,0) je pritom obsiahnutý iba v hrane $e1$ a vrchol (1,1,1) prezmenu iba v hrane $e3$. Preto tieto 2 hrany sú jadrové a tvoria tak jadro u danej funkcie. Nakoľko hrana $e2$ je kompletne pokrytá hranami tvoriacími jadro, môžme túto hranu z pokrytia

vynechať(ako je znázornené na obrázku b)) a z DNF môžme tiež vynechať implikant, ktorý jej prináležal. Dostávame tak Quinovu DNF, ktorá bude mať tvar $D = \bar{x_2}\bar{x_3} \wedge x_1x_3$.

(a) maximálne hrany (b) jadrové hrany

Obr. 3.1: Príklad jadrových hrán ČNBF

Počet hrán, ktoré tvoria jadro grafu G, budeme označovať premennou c_G a $E(c_G)$ bude označovať strednú hodnotu tejto premennej.

3.2.2 Stredná hodnota počtu jadrových hrán

Lema 3.2.1 *Nech* $p_1, p_2, p_3 \in \langle 0, 1\rangle$ *a* $p_1 + p_2 + p_3 = 1$ *potom*

$$E(c_{Gmin}) = \sum \binom{n}{k} 2^{n-k}((p_1 + p_3)^{2^k} - p_3^{2^k})(1 - (1 - p_2^{n-k})^{2^k}) \leq E(c_G)$$
$$\leq \sum \binom{n}{k} 2^{n-k}((p_1+p_3)^{2^k} - p_3^{2^k})(1-(1-(p_2+p_3)^{n-k})^{2^k})(1-p_3(p_1+p_3)^{2^k-1})^{n-k} = E(c_{Gmax}) \tag{3.1}$$

Dôkaz 3.2.1 Počet všetkých k-rozmerných hrán v rámci n-rozmernej kocky B^n je rovný $\binom{n}{k}2^{n-k}$. Kedže chceme, aby táto hrana bola prípustná, potrebujeme, aby táto hrana obsahovala iba vrcholy z množín N_f a $N_{\tilde{f}}$, pričom ale, aspoň jeden z toho musí byť s hodnotou 1. Preto pravdepodobnosť, že hrana bude spĺňať túto vlastnosť je rovná $((p_1 + p_3)^{2^k} - p_3^{2^k})$.

Nech $P_{n,k}$ je pravdepodobnosť, že hrana rozmeru k je jadrovou hranou, potom strednú hodnotu počtu jadrových hrán vieme vyjadriť ako

$$E(c_G) = \sum \binom{n}{k} 2^{n-k}((p_1 + p_3)^{2^k} - p_3^{2^k})P_{n,k}$$

Nakoľko $P_{n,k}$ má rovnakú hodnotu pre všetky hrany rádu k, stačí túto pravdepodobnosť určiť práve pre jednu z nich.

Nech M je nejaká fixovaná hrana rádu k. Potom pravdepodobnosť, že táto hrana je jadrovou hranou, je z definície rovná pravdepodobnosti, že v tejto hrane existuje vrchol nevyskytujúci sa v žiadnej inej maximálnej hrane.

Analýza prípadov nám dáva nasledujúce 2 základné prípady:

(a) prípad, kedy sa vrchol s určitosťou nevyskytuje v žiadnej inej hrane, a teda hrana je jadrová,

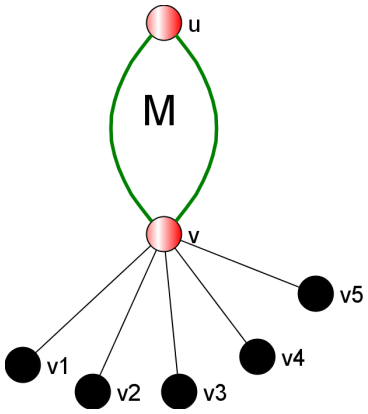
(b) prípad, kedy je vrchol určite obsiahnutý aspoň jednou maximálnou hranou rôznou od M, a teda jeho hrana jadrovou nie je.

Teraz sa bližšie pozrieme na to, kedy nastávajú jednotlivé prípady.

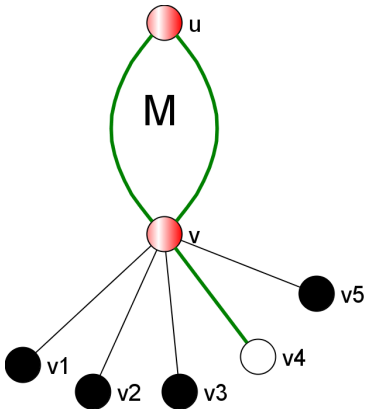
(a) Vrchol v z hrany M sa s určitosťou nevyskytuje v žiadnej inej maximálnej hrane vtedy, keď všetci jeho susedia mimo túto hranu M nadobúdajú hodnotu 0. To z dôvodu, že žiadna hrana nemôže prechádzať cez tieto nulové vrcholy, a tak pokrývať náš vrchol v. (Ilustrované na obrázku 3.2(a))

(b) Vrchol v je s určitosťou súčasťou nejakej inej maximálnej hrany vtedy, keď aspoň 1 jeho sused mimo hrany M nadobúda hodnotu 1. Ak sa tak deje, znamená to, že musí existovať hrana minimálne o dimenzii 1 (pozostávajúca iba z týchto 2 vrcholov), ktorá tieto 2 vrcholy pokrýva.
To, nakoľko sa dá táto hrana potenciálne rozšíriť a akú dimenziu v skutočnosti má, je pre náš ďalší výskum nepodstatné.

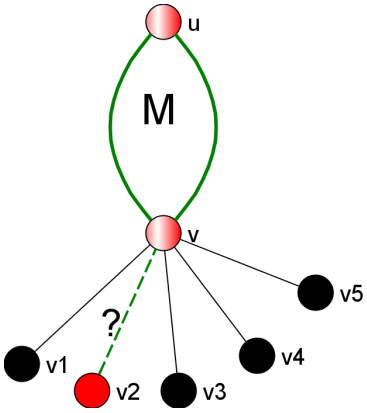
Nastáva však otázka, čo vieme povedať o zvyšnom prípade (3.2(c)), keď aspoň jeden sused vrchola v je nedefinovaný a zároveň žiaden z nich nie je rovný 1?



(a) všetci susedia sú rovní 0



(b) existuje nejaký 1-vý sused



(c) neexistuje žiaden 1-vý sused,
ale existuje aspoň 1 nedefinovaný

Obr. 3.2: Možnosti rozloženia hodnôt susedov mimo hrany M

V takomto prípade, aby sme vedeli s určitosťou povedať, či daný vrchol nie je obsiahnutý žiadnou inou hranou (a tak spôsobuje jadrovosť hrany M), môže vznikať potreba skúmať rozloženie typov susedných vrcholov, a to potenciálne až do vzdialenosti n. Vzhľadom na túto skutočnosť sme sa rozhodlii,

že vyhodnocovanie týchto prípadov obmedzíme na skúmanie najbližších mimohranových susedov úrovne 1. To spôsobí, že vyjadrovaný počet jadrových hrán nebude stanovený presnou hodnotou, ale iba jej dolnou a hornou hranicou.

Nadväzujúc na spomínané prípady, stanovíme postačujúcu podmienku (1.) a nutnú podmienku (2.) toho, aby hrana bola jadrovou. Tieto podmienky zároveň určia dolnú a hornú hranicu počtu jadrových hrán.

Podmienky sú nasledovné:

1. "Hrana M obsahuje vrchol v, ktorého všetci susedia mimo tejto hrany sú rovní 0".
2. "Hrana M obsahuje vrchol v, pre ktorý platí, že každý jeho sused mimo tejto hrany je buď rovný 0, alebo je nedefinovaný".

Teraz tieto podmienky preformulujeme do ekvivalentných znení, ktoré budeme ľahšie vedieť prepísať do matematickej formy.

1. "Neplatí, že pre všetky vrcholy v rámci hrany M existuje aspoň 1 susedný vrchol mimo hrany M, taký že tento vrchol je rovný 1 alebo nie je definovaný".
2. "Neplatí, že pre všetky vrcholy v rámci hrany M existuje aspoň 1 susedný vrchol mimo hrany M, taký že tento vrchol je rovný 1".

Samotným prepísaním týchto podmienok, kde $P1_{n,k}$ prislúcha podmienke (1.) a $P2_{n,k}$ podmienke (2.), dostávame:

$$P1_{n,k} = 1 - (1 - (1 - p_1 - p_3)^{n-k})^{2^k} = 1 - (1 - p_2^{n-k})^{2^k}$$

$$P2_{n,k} = 1 - (1 - (1 - p_1)^{n-k})^{2^k} = 1 - (1 - (p_2 + p_3)^{n-k})^{2^k}$$

Navyše, kedže v prípade jadrových hrán má zmysel uvažovať iba tie, ktoré sú maximálne, musíme vylúčiť tie prípady, kedy sused vrchola v je nedefinovaný a zároveň by sa dala zväčšiť dimenzia hrany. To nastáva s pravdepodobnosťou $p_3(p_1 + p_3)^{2^k-1}$.
Aby to platilo pre všetkých susedov vrchola v, dostávame pravdepodobnosť $(1 - p_3(p_1 + p_3)^{2^k-1})^{n-k}$. Preto $P2_{n,k}$, určujúca hornú hranicu pravdepodobnosti, že hrana bude jadrová má tvar

$$P2_{n,k} = 1 - (1 - (p_2 + p_3)^{n-k})^{2^k}(1 - p_3(p_1 + p_3)^{2^k-1})^{n-k}$$

Kedže $P1_{n,k} \leq P_{n,k} \leq P2_{n,k}$, dostávame z toho, že

$$E(c_{Gmin}) = \sum \binom{n}{k} 2^{n-k}((p_1 + p_3)^{2^k} - p_3^{2^k})(1 - (1 - p_2^{n-k})^{2^k}) \leq E(c_G) \leq$$

$$\leq \sum \binom{n}{k} 2^{n-k}((p_1+p_3)^{2^k}-p_3^{2^k})(1-(1-(p_2+p_3)^{n-k})^{2^k})(1-p_3(p_1+p_3)^{2^k-1})^{n-k} = E(c_{Gmax})$$

čím sme dokázali znenie lemy. □

3.2.3 Odhad počtu jadrových hrán

V tejto časti za pomoci strednej hodnoty počtu jadrových hrán, kombinatorických metód a pravdepodobnostných nerovností asymptoticky ohraničíme počet jadrových hrán obsiahnutých v grafe B^n.

Nasledujúcou lemou určíme hodnotu horného ohraničenia parametra $E(c_G)$.

Lema 3.2.2 *Nech* $\{a_n\}_{n\geq 0}$ *je postupnosť čísel taká, že* $\lim_{n\to\infty} a_n = 0$. *Potom platí*

$$\lim_{n\to\infty} P[E(c_G) \leq n^{(1+a_n)\lg\log_b n}(2p_2)^n] = 1$$

kde $b = 1/(p_1+p_3)$.

Dôkaz 3.2.2 Aby sme dospeli k hodnote tohto ohraničenia, budeme rozoberať sumu $E(c_{Gmax})$ z vety 3.2.1, ktorá je strednou hodnotou parametra vyjadrujúceho maximálny počet jadrových hrán v grafe. Z práce [5] vieme, že B^n s pravdepodobnosťou idúcou k 1, pre $n \to \infty$, neobsahuje žiadnu hranu rozmeru väčšieho ako $\mu = \lfloor \lg n - \lg\lg b \rfloor + 1$. Preto si $E(c_{Gmax})$ rozdelíme na 2 sumy a to také, kde prvá z nich S_1 bude sumou pre $k \lg n + l$ a S_2 bude sumou pre $k > \lg n + l$.

Najprv sa pozrieme na sumu S_2. Ukážeme, že pre vhodne zvolený parameter l a pre dostatočne veľké n, bude S_2 nadobúdať hodnotu 0, a to z dôvodu, že ak B^n neobsahuje žiadnu hranu rozmeru $k > \lfloor \lg n - \lg\lg b \rfloor + 1$, nemôže obsahovať ani hrany rozmeru k, ktoré by boli jadrové.

Konkrétne keďže $0 \leq ((p_1+p_3)^{2^k} - p_3^{2^k}) \leq 1$, $0 \leq (1-(1-(p_2+p_3)^{n-k})^{2^k}) \leq 1$ a $0 \leq (1-p_3(p_1+p_3)^{2^k-1})^{n-k} \leq 1$, tak aj $((p_1+p_3)^{2^k} - p_3^{2^k})(1-(1-(p_2+p_3)^{n-k})^{2^k})(1-p_3(p_1+p_3)^{2^k-1})^{n-k}$ nadobúda hodnotu v intervale $\langle 0,1\rangle$, preto

$$S_2 = \sum_{k>\lg n+l} \binom{n}{k} 2^{n-k}((p_1+p_3)^{2^k} - p_3^{2^k})(1-(1-(p_2+p_3)^{n-k})^{2^k})(1-p_3(p_1+p_3)^{2^k-1})^{n-k}$$

$$\leq \sum_{k>\lg n+l} \binom{n}{k} 2^{n-k} \leq (p_1+p_3)^{2^{\lg n+l}} \sum_{k\geq 0} \binom{n}{k} 2^{n-k}$$

$$= (p_1 + p_3)^{n2^l} 3^n = (3(p_1 + p_3)^{2^l})^n$$

Preto ak zvolíme l tak, aby platilo, že $(p_1 + p_3)^{2^l} < 1/3$, tak $\lim_{n\to\infty} S_2 = 0$.

Druhá časť, suma S_1, ktorá je sumou pre zvyšné hodnoty, a teda pre $k \leq \lg n + l$, teraz určí hľadaný horný odhad $E(c_G)$.

$$S_1 = \sum_{k\leq \lg n+l} \binom{n}{k} 2^{n-k}((p_1+p_3)^{2^k} - p_3^{2^k})(1-(1-(p_2+p_3)^{n-k})^{2^k})(1-p_3(p_1+p_3)^{2^k-1})^{n-k}$$

Kedže $0 \leq (1 - p_3(p_1 + p_3)^{2^k-1})^{n-k} \leq 1$, máme, že

$$S_1 \leq \sum_{k\leq \lg n+l} \binom{n}{k} 2^{n-k}((p_1 + p_3)^{2^k} - p_3^{2^k})(1 - (1 - (p_2 + p_3)^{n-k})^{2^k})$$

Nech $a = -(p_2+p_3)^{n-k}$ a $m = 2^k$. Využijeme teraz Bernoulliho nerovnosť a totiž, že pre ľubovoľné $a > -1$ a $m \geq 0$ platí $(1 - ma) \leq (1 - a)^m$.
Jej použitím toho dostávame

$$S_1 \leq \sum_{k\leq \lg n+l} \binom{n}{k} 2^{n-k}(p_1 + p_3)^{2^k}(1 - (1 - 2^k(p_2 + p_3)^{n-k}))$$

$$= \sum_{k\leq \lg n+l} \binom{n}{k} 2^{n-k}(p_1 + p_3)^{2^k} 2^k (p_2 + p_3)^{n-k}$$

$$= 2^n \sum_{k\leq \lg n+l} \binom{n}{k} (p_1 + p_3)^{2^k} (p_2 + p_3)^{n-k}$$

Aby sme vedeli tento výraz zhora ohraničiť, označme teraz toto naše vnútro sumy ako b_k, teda $b_k = \binom{n}{k}(p_1 + p_3)^{2^k}(p_2 + p_3)^{n-k}$. Potom podiel týchto členov

$$\frac{b_{k+1}}{b_k} = \frac{(n-k)(p_1 + p_3)^{2^k}}{(k+1)(p_2 + p_3)}$$

je menší ako 1 pre $k > \lg\log_{1/(p_1+p_3)} n$ a väčší ako 1 pre $k \leq \lg\log_{1/((p_1+p_3))} n$. Preto dostávame, že maximálna hodnota výrazu b_k je dosahovaná buď pre $k = \lambda$, alebo pre $k = \lambda + 1$, kde

$$\lambda = \lfloor \lg\log_{1/((p_1+p_3))} n \rfloor \tag{3.2}$$

Preto vieme sumu S_1 ohraničiť nasledovne

$$\begin{aligned} S_1 &\leq 2^n(\lg n + l + 1) \max_{k \leq \lg n + l} b_k \\ &\leq 2^n(\lg n + l + 1)\binom{n}{\lambda+1}(p_1+p_3)^{2^\lambda}(p_2+p_3)^{n-(\lambda+1)} \\ &= \binom{n}{\lambda+1}(p_1+p_3)^{2^\lambda}(\lg n + l + 1)2^n(p_2+p_3)^{n-(\lambda+1)} \end{aligned}$$

Využitím nerovnosti $\binom{n}{\lambda+1} < n^{\lambda+1}$ dostávame

$$S_1 \leq n^{\lambda+1}(p_1+p_3)^{2^\lambda}(\lg n + l + 1)2^n(p_2+p_3)^{n-(\lambda+1)}$$

$$= n^{\lambda+1}(p_1+p_3)^{2^\lambda}(\lg n + l + 1)(p_2+p_3)^{-(\lambda+1)}(2(p_2+p_3))^n$$

Zo vzťahu (3.2) vyplýva, že $(p_1+p_3)^{2^\lambda} \leq \frac{1}{n}$, a preto z toho dostávame

$$S_1 \leq n^{\lambda+1}\frac{1}{n}(\lg n + l + 1)(p_2+p_3)^{-(\lambda+1)}(2(p_2+p_3))^n$$

$$\leq n^\lambda(\lg n + c_1)(p_2+p_3)^{-(\lambda+1)}(2(p_2+p_3))^n$$

$$= n^{\lambda} n^{\frac{\lg(\lg n + c_1)}{\lg n}} n^{-\frac{(\lambda+1)\lg(p_2+p_3)}{\lg n}} (2p_2)^n$$

$$= n^{(1+\varepsilon_1(n))\lg\log_{1/((p_1+p_3))} n} (2(p_2+p_3))^n$$

kde $\varepsilon_1(n) = \frac{\lg(\lg n + c_1) - (\lg\log_{1/((p_1+p_3))} n + 1)\lg(p_2+p_3)}{\lg n \lg\log_{1/((p_1+p_3))} n} \to 0$

Toto ohraničenie už udáva znenie lemy, a tým sme ju dokázali. □

V nasledujúcej leme vyjadríme dolný odhad parametra $E(c_G)$.

Lema 3.2.3 *Nech* $\{a_n\}_{n\geq 0}$ *je postupnosť čísel taká, že* $\lim_{n\to\infty} a_n = 0$. *Potom platí*

$$\lim_{n\to\infty} P[E(c_G) \geq n^{(1-a_n)\lg\log_b n}(2p_2)^n] = 1$$

kde $b = 1/(p_1 + p_3)$.

Dôkaz 3.2.3 Budeme postupovať podobným spôsobom ako pri dokazovaní lemy 3.2.2 a sumu udávajúcu strednú hodnotu dolného ohraničenia jadrových hrán $E(c_{Gmax}) = \sum \binom{n}{k} 2^{n-k}((p_1+p_3)^{2^k} - p_3^{2^k})(1-(1-p_2^{n-k})^{2^k})$ si rozdelíme na 2 časti. Už vieme, že B^n pre $n \to \infty$ s pravdepodobnosťou idúcou k 1 neobsahuje žiadnu hranu rozmeru väčšieho ako $\mu = \lfloor \lg n - \lg\lg b \rfloor + 1$, a preto nemôže obsahovať ani žiadnu jadrovú hranu tohto rozmeru. Ak si preto rozdelíme sumu na 2 časti, kde prvá z nich S_2 bude sumou pre $k > \lg n + l$ a S_1 bude sumou pre $k \leq \lg n + l$, stačí preto zvoliť vhodne parameter l spĺňajúci nerovnosť $(p_1+p_3)^{n2^l} < 1/3$ a $\lim_{n\to\infty} S_2 = 0$.

Sumou S_1 pre hodnoty $k \leq \lg n + l$ teraz určíme dolný odhad $E(c_G)$. Aby sme vedeli tento výraz zdola ohraničiť, vypočítame si, pre aké k nadobúda vnútro sumy S_1 maximálne hodnoty.
Po využití Bernoulliho nerovnosti dostávame

$$S_1 \leq \sum_{k \leq \lg n + l} \binom{n}{k} 2^{n-k} (p_1 + p_3)^{2^k} (1 - (1 - 2^k p_2^{n-k}))$$

a teda

$$S_1 \leq 2^n \sum_{k \leq \lg n + l} \binom{n}{k} (p_1 + p_3)^{2^k} p_2^{n-k}$$

Nech $b_k = \binom{n}{k} (p_1 + p_3)^{2^k} p_2^{n-k}$, potom podiel členov

$$\frac{b_{k+1}}{b_k} = \frac{(n-k)(p_1 + p_3)^{2^k}}{(k+1)p_2}$$

je menší ako 1 pre $k > \lg \log_{1/(p_1+p_3)} n$ a väčší ako 1 pre $k \leq \lg \log_{1/((p_1+p_3))} n$. Preto maximálna hodnota výrazu b_k je dosahovaná buď pre $k = \lambda$, alebo pre $k = \lambda + 1$, kde

$$\lambda = \lfloor \lg \log_{1/((p_1+p_3))} n \rfloor \tag{3.3}$$

Keďže $\lambda = \lfloor \lg \log_{1/((p_1+p_3))} n \rfloor \leq \lg n + l$ suma S_1 obsahuje vnútro s $k = \lambda$, a preto vieme sumu S_1 ohraničiť nasledovne

$$S_1 \geq \binom{n}{\lambda} 2^{n-\lambda} (p_1 + p_3)^{2^\lambda} (1 - (1 - p_2^{n-\lambda})^{2^\lambda})$$

Stanovme $a = p_2^{n-k}$ a $m = 2^k$. Využijeme teraz nerovnosť $(1 - ma) \leq (1 - a)^m \leq (1 - ma + m^2 a^2)$, ktorá platí pre $0 \leq a \leq 1$ a $0 \leq m$.
Jej použitím dostávame

$$\geq \binom{n}{\lambda} 2^{n-\lambda} (p_1 + p_3)^{2^\lambda} (1 - (1 - 2^\lambda p_2^{n-\lambda} + (2^\lambda p_2^{n-\lambda})^2))$$

$$= \binom{n}{\lambda} 2^{n-\lambda} (p_1 + p_3)^{2^\lambda} (1 - (1 - 2^\lambda p_2^{n-\lambda} (1 - 2^\lambda p_2^{n-\lambda})))$$

$$= \binom{n}{\lambda} 2^{n-\lambda}(p_1+p_3)^{2^\lambda} 2^\lambda p_2^{n-\lambda}(1-2^\lambda p_2^{n-\lambda})$$
$$= \binom{n}{\lambda} (p_1+p_3)^{2^\lambda} 2^n p_2^{n-\lambda}(1-2^\lambda p_2^{n-\lambda})$$

Keďže $\lim_{n\to\infty} 1-2^\lambda p_2^{n-\lambda} = 1$, pre dostatočne veľké n dostávame

$$S_1 \geq \binom{n}{\lambda} (p_1+p_3)^{2^\lambda} 2^n p_2^{n-\lambda}$$

Použitím nerovnosti $\binom{n}{\lambda} > (\frac{n}{\lambda})^\lambda$ a substitúciou λ dostávame

$$S_1 \geq (\frac{n}{\lambda})^\lambda (p_1+p_3)^{2^\lambda} 2^n p_2^{n-\lambda}$$

$$\geq n^\lambda \lambda^{-\lambda} (p_1+p_3)^{2^\lambda} 2^n p_2^{n-\lambda}$$

Využitím vzťahu (3.3) a substitúciou λ dostávame

$$S_1 \geq n^\lambda \lambda^{-\lambda} n^{-\frac{1}{2}} 2^n p_2^{n-\lambda}$$

$$= n^\lambda n^{-\frac{1}{2}} \lambda^{-\lambda} p_2^{-\lambda} (2p_2)^n$$

$$= n^\lambda n^{-\frac{1}{2}} n^{-\frac{\lambda \log \lambda}{\log n}} n^{-\frac{\lambda \log p_2}{\log n}} (2p_2)^n$$

$$= n^{(1-\varepsilon_2(n)) \lg \log_{1/((p_1+p_3))} n} (2p_2)^n$$

kde $\varepsilon_2(n) = \frac{\frac{1}{2}\lg n + \lg\log_{1/((p_1+p_3))} n \lg\lg\log_{1/((p_1+p_3))} n + \lg\log_{1/((p_1+p_3))} n \lg p_2}{\lg n \lg\log_{1/((p_1+p_3))} n} \to 0$

Toto ohraničenie nám už určuje znenie lemy. □

Na základe ohraničení z predchádzajúcich liem a použitím Markovovej nerovnosti, môžme sformulovať vetu sumarizujúcu jednotlivé ohraničenia počtu jadrových hrán.

Veta 3.2.4 *Počet jadrových hrán c_G obsiahnutých v grafe zodpovedajúcom čiastočnej náhodnej booleovskej funkcii je*

$$n^{(1+o(1))\lg\log\frac{1}{(p_1+p_3)}n}(2p_2)^n \leq c_G \leq n^{(1+o(1))\lg\log\frac{1}{(p_1+p_3)}n}(2(p_2+p_3))^n$$

3.2.4 Zložitosť Quinovej DNF

Na základe určeného počtu jadrových hrán sa môžeme teraz pokúsiť vyjadriť akú zložitosť nadobúda DNF v Quinovom tvare (označovať ju budeme $l(D_Q(f))$). Ako už vieme z definície Quinovej DNF, je to DNF, ktorú dostaneme zo skrátenej DNF vynechaním všetkých tých konjunkcií, ktoré sú pokryté jadrom. Tento počet budeme označovať c_Q.
Aby sme vedeli určiť počet takýchto konjunkcií (hrán im ekvivalentným) z počtu jadrových hrán, ktorý sme získali výpočtom, využijeme niektoré poznatky a úvahy na odhadnutie tohto počtu.
Z práce [5] vieme, že rozmer každej z hrán funkcie f je nanajvýš rovný nejakému m. Preto u každej funkcie f počet bodov prislúchajúcich jadrovým hranám nie je väčší ako počet jadrových hrán(ďalej už len c_G) $\times 2^m$ a zároveň dostávame, že cez každý vrchol prechádza nanajvýš $\sum_{k=0}^{m}\binom{n}{k}$ hrán.
Aby sme ohodnotili počet hrán, ktoré sú pokryté jadrovými hranami, využijeme poznatok, že tento počet nie je väčší od počtu hrán obsahujúcich aspoň jeden vrchol prináležajúci jadrovej hrane. A preto $c_Q \leq c_G 2^m \sum_{k=0}^{m}\binom{n}{k}$.
Keďže v prípade jadrových hrán máme parameter c_G daný horným a spodným ohraničením, pri určovaní zložitosti tejto DNF využijeme iba jeho hornú hranicu a teda počet hrán pokrytých jadrom bude daný výrazom

$$c_Q \leq n^{(1+o(1))\lg\log\frac{1}{(p_1+p_3)}n}(2(p_2+p_3))^n 2^{\lg n+l}\sum_{k=0}^{\lg n+l}\binom{n}{k}$$

Aby sme určili zložitosť Quinovej DNF uvedieme si zložitosť skrátenej DNF, ktorá bola vyjadrená v práci [5] a ktorá je daná vzťahom

$$2^n n^{(1-\epsilon_2(n))\lg\log_{1/(p_1+p_3)} n} \leq l(D_s(f)) \leq 2^n n^{(1+\epsilon_3(n))\lg\log_{1/(p_1+p_3)} n}$$

kde $\epsilon_2(n) = \epsilon_3(n) = O(\frac{1}{\lg\log_{1/(p_1+p_3)} n})$

Keďže

$$l(D_Q(f)) = l(D_s(f)) - c_Q$$

máme

$$l(D_Q(f)) = 2^n n^{(1+o(n))\lg\log_{1/(p_1+p_3)} n} - n^{(1+o(1))\lg\log\frac{1}{(p_1+p_3)}\, n}(2(p_2+p_3))^n 2^{\lg n+l}\sum_{k=0}^{\lg n+l}\binom{n}{k}$$

$$= n^{((1+o(n))\lg\log_{1/(p_1+p_3)} n)} 2^n (1 - (p_2+p_3)^n 2^{\lg n+l}\sum_{k=0}^{\lg n+l}\binom{n}{k})$$

Keďže pre $n \to \infty$ výraz $(p_2+p_3)^n 2^{\lg n+l}\sum_{k=0}^{\lg n+l}\binom{n}{k}$ nadobúda minimálne hodnotytoho dostávame, že na základe našich úvah zložitosť Quinovej DNF $\sim$ zložitosť skrátenej DNF a teda, že uskutočnenie transformácie na Quinovu DNF prinesie iba malú úsporu v ďaľšom hľadaní minimálnej DNF.

3.3 Regulárne vrcholy

Regulárnimi vrcholmi nazývame také vrcholy, ktoré musia byť pokryté v ľubovoľnom vrcholovom pokrytí grafu G.

3.3.1 Zadefinovanie pojmov

Definícia 3.3.1 *Nech graf G je geometrickou reprezentáciou niektorej čiastočnej booleovskej funkcie. Vrchol $u \in V(G)$ nazývame regulárny, ak existuje taký*

vrchol v, $v \neq u$, že pre každú maximálnu hranu K obsiahnutú v G platí $v \in K \Rightarrow u \in K$. Hovoríme, že takýto vrchol v indukuje vrchol u.

Príklad 3.3.1 Majme čiastočnú booleovskú funkciu, ktorá je reprezentovaná hyperkockou s 3 maximálnymi hranami $e1$, $e2$, $e3$ (obrázkok (a)). V rámci tohto grafu je vrchol $u = (1,0,0)$ obsiahnutý množinou hrán $M_u = \{e1, e2\}$, zatiaľ čo vrchol $v = (0,0,0)$ iba hranami $M_v = \{e1\}$. Keďže $M_v \subseteq M_u$, tak vrchol v indukuje vrchol u, ktorý je tým pádom regulárnym vrcholom. Obdobným spôsobom sa dá ukázať regulárnosť vrchola so súradnicami $(1,0,1)$ (obrázkok (b)).
Nakoľko skúmanej funkcii prislúcha skrátená DNF $D = \bar{x_2}\bar{x_3} \wedge x_1\bar{x_2} \wedge x_1x_3$ a hrana $e2$ je regulárna (skladá sa čisto z regulárnych vrcholov), môžme túto hranu vynechať z pokrytia grafu a zo skrátenej DNF môžme tiež vynechať implikant, ktorý jej prináležal. Dostávame tak DNF typu $\sum T$, ktorá má tvar $D = \bar{x_2}\bar{x_3} \wedge x_1x_3$.

(a) maximálne hrany

(b) regulárne vrcholy

Obr. 3.3: Príklad regulárnych vrcholov ČNBF

Počet vrcholov, ktoré v grafe G spĺňajú podmienku regulárnosti budeme označovať premennou r_G a jej strednú hodnotu budeme vyjadrovať ako $E(r_G)$.

3.3.2 Stredná hodnota počtu regulárnych vrcholov

Lema 3.3.1 *Nech $p_1, p_2, p_3 \in \langle 0,1 \rangle$ a $p_1 + p_2 + p_3 = 1$ potom*

$$E(r_G) \leq 2^n \sum_{k=1}^{n} \binom{n}{k} ((p_1+p_3)^{2^k} - p_3^{2^k})((p_2+p_3) + p_1(p_1+p_3)^{2^k-1})^{n-k}$$

Dôkaz 3.3.1 $P_n(u)$ nech je pravdepodobnosť, že vrchol u je regulárny vrchol.
Kedže takáto pravdepodobnosť, že vrchol je regulárny sa vyjadruje rovnako pre každý z vrcholov B^n, dostávame

$$E(r_G) = 2^n P_n(u)$$

Označením $P_{n,v}(u)$ budeme udávať pravdepodobnosť, že vrchol u je regulárny vrchol a je indukovaný vrcholom v. Táto pravdepodobnosť je vo všeobecnosti nemenná pre všetky tie vrcholy $v_1, v_2, \cdots, v_m$, ktoré majú od vrchola u rovnakú vzdialenosť.

Nech $P_{n,k}$ značí tú pravdepodobnosť, že vrchol v indukuje vrchol u a jeho vzdialenosť je rovná k. Potom platí

$$P_n(u) \leq \sum_{k=1}^{n} \binom{n}{k} P_{n,k}(u)$$

Nerovnosť dostávame práve preto, že jeden vrchol môže byť indukovaný naraz viacerými rôznymi vrcholmi.

Uvažujme preto teraz vrcholy u a v vzdialené od seba vzdialenosťou k. Zadefinujme najmenšiu možnú hranu, ktorá bude oba z nich obsahovať. (Kedže vrchol v indukuje vrchol u, musí existovať aspoň 1 taká hrana, ktorá obsahuje oba vrcholy. Budeme uvažovať najmenšiu z nich.) Kedže naše vrcholy sú vo vzdialenosti k, znamená to, že $n-k$ súradníc majú spoločných. Táto hrana bude mať teda dimenziu tiež k. Budeme ju označovať M.
Nech w je taký vrchol z B^n, že w je susedom s u, ale zároveň tento vrchol nie je súčasťou hrany M. Tento vrchol w sa z definície o susedstve v grafe líši od

vrchola u práve v jednej súradnici. Nech je to súradnica i. Zadefinujeme teraz taký vrchol w', ktorý je susedom vrchola v a líši sa od neho práve v tejto i-tej súradnici. Inak povedané, oba z vrcholov w, w' sú susedmi v jednom a tom istom smere.

Nakoľko pracujeme aj s prípadmi, keď vrcholy w a w' môžu byť nedefinované, nevieme povedať, či vrchol u tvorí s w resp. v s w' nejakú hranu. Preto, aby sme sa vyhli komplikovanému skúmaniu vrcholov do väčšej hĺbky, budeme pri vyčíslovaní pravdepodobnosti $P_{n,k}(u)$ pracovať iba s hornou hranicou tohto počtu(dolnú hranicu nemá význam skúmať, keďže jeden vrchol môže byť indukovaný naraz viacerými inými a to nám spôsobuje, že $P_n(u) \leq \sum_{k=1}^{n} \binom{n}{k} P_{n,k}(u)$).

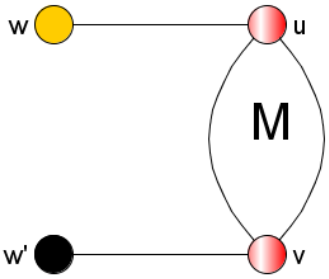
Analýza prípadov nám hovorí, že vrchol v indukuje vrchol u práve vtedy, keď sa udeje jedna z nasledovných možností:

- vrchol w' je rovný 0; vrchol w môže v takom prípade nadobúdať ľubovoľnú hodnotu: toto sa nám vyskytuje s pravdepodobnosťou $p_2(p_1 + p_2 + p_3)$ (obrázok 3.4(a))

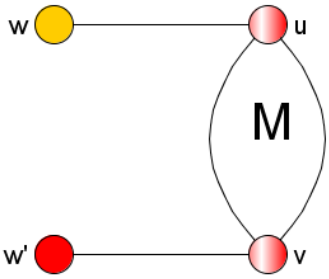
- vrchol w' nie je definovaný; vtedy, ak hľadáme horné ohraničenie počtu regulárnych vrcholov, predpokladáme, že w' nevytvára s v hranu (čo by nás v počte obmedzilo), a teda opäť, w môže nadobúdať ľubovoľnú hodnotu: to dostávame s pravdepodobnosťou $p_3(p_1 + p_2 + p_3)$ (obrázok 3.4(b))

- vrchol w' je rovný 1; je zrejmé, že w' spolu s v sú obsiahnuté v jednej a tej istej hrane. Tým pádom na to, aby v indukoval u, je nevyhnutné, aby aj u bol súčasťou tejto hrany. To nastáva iba vtedy, ak w' s w tvoria hranu o rozmere k, ktorá tak rozširuje hranu M s dimenziou k

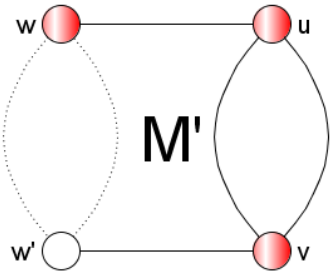
na hranu M' s dimenziou $k+1$: táto možnosť nastáva s pravdepodobnosťou $p_1(p_1+p_3)^{2^k-1}$ (obrázok 3.4(c))



(a) w' je rovné 0 a tak w môže nadobúdať ľubovolné hodnoty



(b) w' je nedefinované a tak w môže byť s ľubovolnou hodnotou



(c) w' je rovné 1 a tak w musí s w' tvoriť hranu o dimenzii k, čím vieme hranu M rozšíriť na M' s dimenziou $k+1$

Obr. 3.4: Možnosti indukovania vrchola u vrcholom v

Preto

$$P_{n,k}(u) \leq ((p_1+p_3)^{2^k} - p_3^{2^k})(p_2(p_1+p_2+p_3)+p_3(p_1+p_2+p_3)+p_1(p_1+p_3)^{2^k-1}))^{n-k}$$

$$\leq ((p_1+p_3)^{2^k} - p_3^{2^k})((p_2+p_3)+p_1(p_1+p_3)^{2^k-1})^{n-k}$$

z toho dostávame

$$P_n(u) \leq \sum_{k=1}^{n} \binom{n}{k} ((p_1+p_3)^{2^k} - p_3^{2^k})((p_2+p_3) + p_1(p_1+p_3)^{2^k-1})^{n-k}$$

a z toho máme

$$E(r_G) \leq 2^n \sum_{k=1}^{n} ((p_1+p_3)^{2^k} - p_3^{2^k})((p_2+p_3) + p_1(p_1+p_3)^{2^k-1})^{n-k}$$

čím máme lemu dokázanú. □

3.3.3 Odhad počtu regulárnych vrcholov

V tejto časti za pomoci strednej hodnoty počtu regulárnych vrcholov, kombinatorických metód a pravdepodobnostných nerovností odhadneme počet regulárnych vrcholov obsiahnutých v grafe B^n.

Veta 3.3.2 *Počet regulárnych vrcholov r_G v grafe G zodpovedajúcom čiastočnej náhodnej booleovskej funkcii je*

$$r_G < \rho^n$$

pričom ρ je konštanta závislá od pravdepodobností p_1 a p_3 s hodnotaou v intervale $1.57 \leq \rho < 2$.

Dôkaz 3.3.2 Stredná hodnota $E(r_G)$ určená predchádzajúcou vetou môže byť ohraničená nasledujúco:

$$E(r_G) \leq 2^n \sum_{k=1}^{n} \binom{n}{k} ((p_1+p_3)^{2^k} - p_3^{2^k})((p_2+p_3) + (p_1+p_3)^{2^k})^{n-k}$$

$$\leq 2^n \sum_{k=1}^{n} \binom{n}{k} (p_1+p_3)^{2^k} ((p_2+p_3)+(p_1+p_3)^{2^k})^{n-k}$$

$$= \sum_{k=1}^{n} \binom{n}{k} 2^k (p_1+p_3)^{2^k} (2((p_2+p_3)+(p_1+p_3)^{2^k}))^{n-k}$$

$$\leq \sum_{k=1}^{n} \binom{n}{k} 2^k (p_1+p_3)^{2k} (2((p_2+p_3)+(p_1+p_3)^{2^k}))^{n-k}$$

$$= \sum_{k=1}^{n} \binom{n}{k} (2(p_1+p_3)^2)^k (2((p_2+p_3)+(p_1+p_3)^{2^k}))^{n-k}$$

$$\leq (2n(p_1+p_3)^2)((p_2+p_3)+(p_1+p_3)^2)^{n-1} + \sum_{k=2}^{n} \binom{n}{k} (2(p_1+p_3)^2)^k 2(((p_2+p_3)+(p_1+p_3)^4))^{n-k}$$

$$\leq (2n(p_1+p_3)^2)((p_2+p_3)+(p_1+p_3)^2)^{n-1} + \sum_{k=0}^{n} \binom{n}{k} (2(p_1+p_3)^2)^k 2(((p_2+p_3)+(p_1+p_3)^4))^{n-k}$$

$$= (2n(p_1+p_3)^2)((p_2+p_3)+(p_1+p_3)^2)^{n-1} + (2((p_2+p_3)+(p_1+p_3)^2+(p_1+p_3))^4)^n$$

$$\leq c_1 \rho^n$$

kde c_1 je konštantou a ρ je konštanta, ktorej hodnoty závisia od určených pravdepodobností p_1, p_2, p_3. Nech f je funkcia o 2 premenných zadefinovaná nasledovne, $f(p_1, p_3) = 2((1-p_1)+(p_1+p_3)^2+(p_1+p_3))^4)$. Táto funkcia na intervale, kde $p_1 \in [0,1]$ a $p_3 \in [0,1]$ nadobúda minimum dané hodnotou 1.5704. Použitím Markovovej nerovnosti dostávame, že $r_G < \rho^n$, čím sme dokázali vetu. □

3.3.4 Regulárne hrany

Aby sme sa mohli zaoberať zložitosťou DNF typu $\sum T$ je pre nás dôležité vyjadriť počet regulárnych hrán v grafe G.

3.3.5 Zložitosť DNF typu $\sum T$

Na základe určeného počtu regulárnych vrcholov môžme teraz pristúpiť k vyjadreniu zložitosti, ktorú nadobúda DNF typu $\sum T$ (označovať ju budeme $l(D_{\sum T}(f))$).
Z publikácie [6] je nám známe tvrdenie o nutnej a postačujúcej podmienke týkajucej sa DNF typu $\sum T$: nutná a postačujúca podmienka na to, aby konjunkcia K zo skrátenej DNF $D_s(f)$ nepatrila do DNF typu $\sum T$ je, aby každý vrchol zodpovedajúcej hrany bol regulárny vzhľadom na graf zodpovedajúci funkcii f.
Aby sme určili počet regulárnych hrán, využijeme fakt, že z práce [5] vieme, že rozmer každej z hrán funkcie f je nanajvýš rovný nejakému m a že tento počet nie je väčší od počtu maximálnych hrán obsahujúcich aspoň jeden regulárny vrchol. Tých je $r_G \sum_{k=0}^{m} \binom{n}{k}$.

Zložitosť DNF typu $\sum T$ je preto daná vzťahom

$$l(D_{\sum T}(f)) = l(D_s(f)) - r_G \sum_{k=0}^{m} \binom{n}{k}$$

Keďže zložitosť skrátenej DNF, ktorá bola vyjadrená v práci [5] je daná vzťahom $2^n n^{(1-\epsilon_2(n)) \lg \log_{1/(p_1+p_3)} n} \leq l(D_s(f)) \leq 2^n n^{(1+\epsilon_3(n)) \lg \log_{1/(p_1+p_3)} n}$, kde $\epsilon_2(n) = \epsilon_3(n) = O(\frac{1}{\lg \log_{1/(p_1+p_3)} n})$

dostávame

$$l(D_{\sum T}(f)) = 2^n n^{(1+\epsilon_3(n)) \lg \log_{1/(p_1+p_3)} n} - 1.57^n \sum_{k=0}^{\lg n+l} \binom{n}{k}$$

Z toho vyplýva, že $l(D_{\sum T}(f)) \sim l(D_s(f))$ a teda aj $l(D_{\sum T}(f)) \sim l(D_Q(f))$.

Kapitola 4

Záver

Cieľom tejto diplomovej práce bolo asymptoticky odhadnúť a ohraničiť počet jadrových hrán a regulárnych vrcholov u čiastočných booleovských funkcií a vyjadriť a porovnať zložitosti Quinovej DNF a DNF typu $\sum T$ so zložitosťou skrátenej DNF.

Pôvodná predstava, že skúmané vlastnosti (počet jadrových hrán, počet regulárnych vrcholov) budú presne vyčíslené sa narušila tým, že definícia prípustnej hrany bola vymedzená tak, že aspoň 1 z jej vrcholov musí byť rovný 1. To spôsobilo, že pri určovaní jadrovosti hrany sme boli nútení skúmať rozloženie hodnôt susedných vrcholov potenciálne až do vzdialenosti n. Preto sme sa rozhodli od určenia presného vyjadrenia počtu upustiť a pokúsili sme sa vyjadriť hornú resp. dolnú hranicu počtu na základe susedov hrany so vzdialenosťou 1. Vychádzajúc zo spomenutého sme dospeli k záveru, že počet jadrových hrán je pre $n \to \infty$ daný rozmedzím

$$n^{(1+o(1))\lg\lg\frac{1}{(p_1+p_3)}\,n}(2p_2)^n \leq c_G \leq n^{(1+o(1))\lg\lg\frac{1}{(p_1+p_3)}\,n}(2(p_2+p_3))^n$$

Pri vyjadrovaní počtu regulárnych vrcholov bolo určujúce, že jeden vrchol môže byť naraz indukovaný viacerými inými vrcholmi. Z toho dôvodu sme mohli určiť iba hornú hranicu tohto počtu, ktorá je daná vzťahom $r_G < \rho^n$,

kde ρ je konštanta závislá od hodnôt p_1 a p_3 a jej hodnota je v rozmedzí $1.57 \leq \rho < 2$.
Tieto vyjadrenia počtov sme následne aplikovali na stanovenie zložitosti predmetných DNF. Quinova DNF po úvahách, koľko hrán je pokrytých jadrom bola určená hornou hranicou a asymptoticky bola rovná zložitosti skrátenej DNF.
Na vyjadrenie zložitosti DNF typu $\sum T$ sme potrebovali určiť počet regulárnych hrán z počtu regulárnych vrcholov. Tento počet bol úvahami ohraničený a aplikovaný na zložitosť DNF, ktorá sa opäť ukázala asymptoticky rovná $l(D_s(f))$.
Ukázali sme teda, že pomer medzi skrátenou a Quinovou DNF, ako aj pomer medzi skrátenou DNF a DNF typu $\sum T$ je asymptoticky rovný a teda zlepšenie pri procese minimalizácie je pri $n \to \infty$ zanedbatelne malé.

Uvedomujeme si, že v problematike čiastočných náhodných booleovských funkcií riešenej v našej diplomovej práci existujú ešte miesta, ktoré majú svoje nedostatky (napr. presnejšie vyjadrenie zložitosti Quinovej DNF a zložitosti DNF typu $\sum T$), preto ich pre ďalší výskum odporúčame podrobnejšie preskúmať.

Literatúra

[1] B. Bollobas. *Random Graphs*. Cambridge University Press, 2001.

[2] Reinhard Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, Heidelberg, third edition, 2005.

[3] Paul Erdos and Joel H Spencer. *Probabilistic methods in combinatorics, by Paul Erdos and Joel Spencer*. Academic Press, New York,, 1974.

[4] Radoslav Harman. Pravdepodobnosť a štatistika. Poznámky k prednáškam, FMFI UK 2007.

[5] Lucia Haviarová. Metrické vlastnosti čiastočných boolovských funkcií. Master's thesis, 2011.

[6] S. V. Jablonskij. *Úvod do diskrétnej matematiky*. ALFA, 1984.

[7] Eduard Toman. Enumerácia diskrétnych štruktúr. Poznámky k prednáškam, FMFI UK 2007.

[8] Eduard Toman and Martin Stanek. Analysis of greedy algorithm for vertex covering of random graph by cubes. *Computers and Artificial Intelligence*, 25(5):393–404, 2006.

[9] Eduard Toman and Martin Stanek. On the number of regular vertices and kernel subgraphs in random graphs. *Journal of Applied Mathematics, Statistics and Informatics*, 1:57–66, 2006.

[10] Eduard Toman, Ľuboš Baník, and Martin Stanek. On structural properties of random subgraphs of n-cube. *Journal of Applied Mathematics, Statistics and Informatics*, 1:71–83, 2007.