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

Metrické vlastnosti čiastočných boolovských funkcií

(Diplomová práca)

Bc. Lucia Haviarová

Vedúci: doc. RNDr. Eduard Toman, CSc.
Odbor: 9.2.1 Informatika
Kód práce: 2dc72081-65ae-4e80-a0df-c695a56a1d8d Bratislava, 2011

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Lucia Haviarová
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 : Metrické vlastnosti čiastočných booleovských funkcií

Cieľ : Cieľom diplomovej práce je získať hodnoty takých parametrov ako sú dĺžka skrátenej dnf., minimálnej a najkratšej dnf., prípadne iredundantnej dnf.

Vedúci : doc. RNDr. Eduard Toman, CSc.

Dátum zadania: 15.11.2009

Dátum schválenia: 18.02.2011

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

Haviarová

študent

vedúci práce

Dátum potvrdenia finálnej verzie práce, súhlas s jej odovzdaním (vrátane spôsobu sprístupnenia)

vedúci práce

Čestne prehlasujem, že som túto diplomovú prácu vypracovala samostatne s použitím citovaných zdrojov a literatúry.

Bratislava, dňa Lucia Haviarová

Poďakovanie

Ďakujem vedúcemu diplomovej práce doc. RNDr. Eduardovi Tomanovi, CSc. za odbornú pomoc, za výber témy, študijné materiály, konzultácie a ostatnú pomoc pri vypracovaní. Ďalej by som chcela poďakovať svojim rodičom za výraznú podporu pri štúdiu a svojmu priateľovi za všetko, čo pre mňa urobil.

Abstrakt

Autor: Lucia Haviarová
Názov diplomovej práce: Metrické vlastnosti čiastočných boolovských funkcií
Škola: Univerzita Komenského v Bratislave
Fakulta: Fakulta matematiky, fyziky a informatiky
Katedra: Katedra informatiky
Vedúci diplomovej práce: doc. RNDr. Eduard Toman, CSc.
Rozsah práce: 56 strán

Bratislava, máj 2011

Cieľom tejto diplomovej práce je sformulovať a dokázať tvrdenia o asymptotických odhadoch zložitosti úplnej a skrátenej disjunktívnej normálnej formy náhodnej čiastočnej boolovskej funkcie, pracovať s geometrickou reprezentáciou tejto funkcie a preskúmať dimenzie a počet podkociek.

KĽÚČOVÉ SLOVÁ: náhodná čiastočná boolovská funkcia, disjunktívna normálna forma, pravdepodobnostný priestor, interval

Abstract

Author: Lucia Haviarová
Name of Diploma Thesis: Metric Properties of Partial Boolean Functions
University: Comenius University, Bratislava, Slovakia
Faculty: Faculty of Mathematics, Physics and Informatics
Department: Department of Informatics
Thesis advisor: doc. RNDr. Eduard Toman, CSc.
Number of Pages: 56

Bratislava, May 2011

The aim of this Diploma Thesis is to formulate and prove the propositions about the asymptotic estimate of a complexity of the principal and the abbreviated disjunctive normal form of a random partial Boolean function, to work with the geometric interpretation of this function and investigate dimensions and a number of subintervals.

KEYWORDS: Random Partial Boolean Function, Disjunctive Normal Form, Probability Space, Interval

Obsah

Zoznam obrázkov

Úvod

Práca je venovaná metrickým vlastnostiam čiastočných boolovských funkcií. Uvedená problematika patrí do oblasti teoretickej informatiky, v ktorej našli už mnoho aplikácií výsledky teórie náhodných grafov. Jednoducho povedané, spravidla nás zaujíma práve priemerný prípad pri použití niektorého algoritmu. Ak sú naším modelom grafy, potom teória náhodných grafov zavádza do týchto modelov pravdepodobnosť.

Boolovské funkcie sú vhodným modelom pri riešení mnohých problémov v rôznych oblastiach. Zohrávajú dôležitú úlohu v otázkach komplexnosti, v návrhu logických obvodov, kde zložitosť schémy priamo súvisí so zložitosťou jemu príslušnej DNF. Vlastnosti boolovských funkcií majú tiež významnú úlohu v kryptológii, hlavne pri návrhu šifrovacích algoritmov so symetrickými klúčmi. Ďalej v algoritmoch pri rozpoznávaní obrazcov a v teórii testov. Oblasť náhodných boolovských funkcií je pomerne dobre preskúmaná, zaoberali sa ňou napríklad Glagolev, Sapozhenko (pravdepodobnostný model s prípadom, keď n-tica sa zobrazí na hodnotu 0 resp. 1 s pravdepodobnosťou $\frac{1}{2}$), ale aj Weber a Yablonski (minimalizácia DNF). V tejto práci sme nadväzovali hlavne na výsledky Škovieru, ktorý skúmal náhodné úplné boolovské funkcie a pracoval so všeobecným modelom pravdepodobnosti.

V úvode sme definovali základné pojmy a zaviedli označenia, ktoré sme v práci používali. Tiež sme definovali minimálnu, najkratšiu, iredundantnú a skrátenú DNF. Formulovali sme úlohu v geometrickej forme, keďže to bol

prístup, akým sme sa na problém pozerali. Ďalej sme uviedli základné pojmy z pravdepodobnosti, ktoré sme používali. V ďalšej časti sme popísali vzťahy medzi parametrami DNF úplných a čiastočných boolovských funkcií a uviedli sme aj významné výsledky niektorých autorov. V ďalších kapitolách sa nachádzajú vlastné výsledky. Pozostávajú zo zadefinovania pravdepodobnostného priestoru, v ktorom sme pracovali. Určili sme strednú hodnotu a disperziu počtu k-rozmerných intervalov funkcie, ohraničili a asymptoticky odhadli ich počet. Na základe takto získaných výsledkov sme mohli odhadnúť zložitosť úplnej DNF čiastočnej boolovskej funkcie. Ďalej sme sa snažili odhadnúť zložitosť skrátenej DNF, ktorá je rovná počtu všetkých maximálnych intervalov. Na to sme potrebovali vypočítať strednú hodnotu počtu k-rozmerných maximálnych intervalov funkcie.

Pri dokazovaní týchto výsledkov sme používali rôzne pravdepodobnostné metódy a kombinatorické odhady, Markovovu a Čebyševovu nerovnosť, strednú hodnotu a disperziu.

Kapitola 1

Základné pojmy a označenia

V úvodnej kapitole zavedieme základné pojmy a označenia ktoré budeme v práci používať. Tiež definujeme minimálnu, najkratšiu, iredundantnú a skrátenú DNF a na ukážku uvedieme niekoľko príkladov. Ďalej formulujeme geometrický prístup k problému, ktorý uprednostníme pred algebraickým a popíšeme pravdepodobnostné metódy, ktoré budeme neskôr potrebovať. Pojmy, definície a výsledky nachádzajúce sa v tejto kapitole sú prebraté z použitej literatúry.

Čiastočná boolovská funkcia premenných $x_1, x_2, \ldots, x_n$ predstavuje zobrazenie, ktoré podmnožine množiny $\{0,1\}^n$ priraďuje hodnotu z množiny $\{0,1\}$. Funkcia teda definuje zobrazenie, ktoré nie je nutne totálne (funkcia nemusí byť definovaná vo všetkých bodoch definičného oboru).

Množinu všetkých $n-$tíc môžeme rozdeliť do troch množín. Prvá množina sú body, v ktorých funkcia nadobúda hodnotu 1, druhá body, v ktorých nadobúda hodnotu 0 a tretia body, v ktorých funkcia nie je definovaná. Existuje teda 3^{2^n} čiastočných boolovských funkcií.

Každú čiastočnú boolovskú funkciu možno realizovať pomocou disjunktívnej normálnej formy (DNF). Ľubovoľná čiastočná boolovská funkcia môže byť vo všeobecnosti vyjadrená vo forme DNF viacerými spôsobmi. V súvislosti s tým vzniká možnosť výberu najvhodnejšej realizácie. Úloha, ktorá sa zaoberá tým, ako pre ľubovoľnú čiastočnú boolovskú funkciu zostrojiť minimálnu DNF vzhľadom na niektorý index jednoduchosti[1], sa nazýva problém minimalizácie čiastočných boolovských funkcií.

Čiastočná náhodná boolovská funkcia premenných $x_1, x_2, \ldots, x_n$ predstavuje zobrazenie, ktorá množine $\{0,1\}^n$ priraďuje hodnotu z množiny $\{0,1\}$ alebo je v danom bode nedefinovaná, v závislosti od parametrov $p_1, p_2, p_3 \in \langle 0,1\rangle$, $p_1 + p_2 + p_3 = 1$[2]. Tieto parametre môžu byť konštantné, alebo môžu závisieť od n. Význam parametrov je taký, že sa každá n-tica zobrazí na hodnotu 1 s pravdepodobnosťou p_1, na hodnotu 0 s pravdepodobnosťou p_2 a nie je v danom bode definovaná s pravdepodobnosťou p_3.

1.1 Minimálna DNF, najkratšia DNF

Označenie 1.1.1

$$x^\sigma = x\sigma \vee \bar{x}\bar{\sigma}$$

kde σ je parameter rovnajúci sa 0 alebo 1. Je zrejmé, že

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

Vidieť, že $x^\sigma = 1$ vtedy a len vtedy, keď $x = \sigma$, t.j. hodnota základu sa rovná hodnote exponentu.

[1] Definované neskôr.

[2] Prípady, keď niektorý z parametrov nadobúda hodnotu 0 alebo 1 sú triviálne, preto ich nebudeme v tejto diplomovej práci uvažovať.

Definícia 1.1.1 *Výraz*

$$\mathrm{K} = x_{i_1}^{\sigma_1} \wedge \ldots \wedge x_{i_r}^{\sigma_r}, \ \ (i_v \neq i_u \ \ pre \ \ v \neq u)$$

sa nazýva elementárnou konjunkciou. Číslo r sa nazýva rádom elementárnej konjunkcie. Definitoricky považujeme konštantu 1 za konjunkciu rádu 0.

Definícia 1.1.2 *Implikantom čiastočnej boolovskej funkcie f nazývame elementárnu konjunkciu K, takú že existuje* $\tilde{\alpha}, f(\tilde{\alpha}) = 1$, *že* $K(\tilde{\alpha}) = 1$ *a zároveň, pre každé* $\tilde{\beta}, f(\tilde{\beta}) = 0$ *platí* $K(\tilde{\beta}) = 0$.

Definícia 1.1.3 *Výraz*

$$\mathrm{D} = \bigvee_{i=1}^{s} \mathrm{K}_i \ \ (\mathrm{K}_i \neq \mathrm{K}_j \ \ pre \ \ i \neq j)$$

kde K_i $(i = 1, \ldots, s)$ *je elementárna konjunkcia rádu* r_i, *sa nazýva disjunktívna normálna forma.*

Pre každú čiastočnú boolovskú funkciu existuje viacero DNF, ktoré ju realizujú. Preto zavedieme index jednoduchosti L(D), ktorý charakterizuje zložitosť DNF D. Pre funkcionál L(D) vyžadujeme splnenie týchto axióm:

I. Axióma nezápornosti. Pre ľubovoľnú DNF $\mathrm{L(D)} \geq 0$.

II. Axióma monotónnosti (vzhľadom na násobenie). Nech $\mathrm{D} = \mathrm{D}' \vee x_i^{\sigma_i}\mathrm{K}'$. Potom $\mathrm{L(D)} \geq \mathrm{L(D}' \vee \mathrm{K}')$.

III. Axióma vypuklosti (vzhľadom na sumáciu). Nech $\mathrm{D} = \mathrm{D}_1 \vee \mathrm{D}_2$. Ak $\mathrm{D}_1 \wedge \mathrm{D}_2 \equiv 0$, tak platí $\mathrm{L(D)} \geq \mathrm{L(D}_1) + \mathrm{L(D}_2)$.

IV. Axióma invariantnosti (vzhľadom na izomorfizmus). Nech DNF D′ bola získaná z DNF D premenovaním premenných (bez stotožnenia). Potom $\mathrm{L(D)} = \mathrm{L(D}')$.

Príklady rozličných indexov jednoduchosti DNF:

1. $l(D)$ je počet symbolov premenných, ktoré sa vyskytujú v zápise DNF D.

2. $L(D)$ je počet elementárnych konjunkcií vyskytujúcich sa v DNF D.

3. $L^{-}(D)$ je počet symbolov s negáciou vyskytujúcich sa v zápise DNF D.

Definícia 1.1.4 *Minimálna DNF D, ktorá realizuje čiastočnú boolovskú funkciu $f(\alpha_1, \ldots, \alpha_n)$ vzhľadom na index jednoduchosti $l(N)$, sa nazýva jednoducho minimálna DNF. Označujeme ju $M(f)$. Minimálna DNF vzhľadom na index jednoduchosti $L(D)$ sa nazýva najkratšou DNF. Označujeme ju $K(f)$.*

1.2 Iredundantná DNF

Nech D je ľubovoľná DNF a

$$\mathrm{D} = \mathrm{D}' \vee \mathrm{K} \qquad a \qquad \mathrm{D} = \mathrm{D}' \vee x_i^{\sigma_i}\mathrm{K}'$$

kde K je nejaká elementárna konjunkcia z D, D′ je DNF vytvorená z ostatných konjunkcií nachádzajúcich sa v D, $x_i^{\sigma_i}$ je neurčitý činiteľ z K a K′ je súčin zvyšných činiteľov z K. Poznáme dva typy transformácií na D:

I. Operácia vynechania elementárnej konjunkcie. Prechod od DNF D k DNF D′ sa nazýva transformácia, ktorá spočíva vo vynechaní elementárnej konjunkcie K. Táto transformácia je definovaná vtedy a len vtedy, keď D = D′.

II. Operácia vynechania činiteľa. Prechodom od DNF D k DNF D′ ∨ K′ je transformácia, ktorá spočíva vo vynechaní činiteľa $x_i^{\sigma_i}$. Táto transformácia je definovaná vtedy a len vtedy, keď D′ ∨ K′ = D.

Definícia 1.2.1 *Prostým implikantom disjunktívnej normálnej formy D čiastočnej boolovskej funkcie budeme nazývať implikant K, na ktorý sa nedá aplikovať transformácia II.*

Definícia 1.2.2 *DNF* D*, ktorú nemožno zjednodušiť pomocou transformácií* 1 *a* 2 *sa nazýva iredundantnou DNF čiastočnej boolovskej funkcie f vzhľadom na transformácie I a II. Označujeme ju* $T(f)$.

1.3 Formulácia úlohy v geometrickej forme

V tejto práci budeme preferovať geometrický prístup k minimalizačnému problému čiastočnej boolovskej funkcie pred algebraickým. Množinu všetkých čiastočných boolovských funkcií n premenných budeme označovať ako P^n a množinu všade definovaných funkcií budeme označovať ako $\tilde{P}^n$. Množinu všetkých binárnych súborov dĺžky n nazývame jednotková n-rozmerná kocka B^n alebo n-rozmerná kocka B^n. Váhou súboru $(\alpha_1, \ldots, \alpha_n)$ nazývame súčet jeho súradníc. B_k^n označujeme množinu súborov váhy k z B^n.

Označenie 1.3.1 *Nech* $f(\alpha_1, \ldots, \alpha_n)$ *je ľubovoľná náhodná čiastočná boolovská funkcia. Pre túto funkciu zadefinujeme tri podmnožiny vrcholov kocky* B^n *nasledovne*

$$(\alpha_1, \ldots, \alpha_n) \in N_f \iff f(\alpha_1, \ldots, \alpha_n) = 1$$

$$(\alpha_1, \ldots, \alpha_n) \in N_f^- \iff f(\alpha_1, \ldots, \alpha_n) = 0$$

$$(\alpha_1, \ldots, \alpha_n) \in N_{\tilde{f}} \iff \textit{funkcia f nie je definovaná v bode } (\alpha_1, \ldots, \alpha_n)$$

Vrcholy množiny N_f budeme na obrázkoch znázorňovať bielou farbou, vrcholy množiny $N_{\tilde{f}}$ budeme znázorňovať červenou farbou a vrcholy množiny N_f^- čiernou.

Z ľubovoľnej dvojice množín N_f, $N_{\tilde{f}}$, N_f^- sa pôvodná čiastočná boolovská funkcia f určuje spätne jednoznačne.

N-rozmernú kocku B^n môžeme jednoducho rozšíriť na graf takým spôsobom, že dva vrcholy $\alpha,\ \beta \in B^n$ budú susedné práve vtedy, keď sú odlišné práve v jednej súradnici. Teda B^n má 2^n vrcholov a $n * 2^{n-1}$ hrán.

Definícia 1.3.1 *Nech $\sigma_{i_1}, \ldots, \sigma_{i_r}$ je pevne zvolená r-tica čísel z 0 a 1 taká, že $1 \leq i_1 < \ldots < i_r \leq n$. Množina všetkých vrcholov $(\alpha_1, \ldots, \alpha_n)$ kocky B^n takých, že $\alpha_{i_1} = \sigma_{i_1}, \alpha_{i_2} = \sigma_{i_2}, \ldots, \alpha_{i_r} = \sigma_{i_r}$ sa nazýva (n-r)-rozmernou hranou.*

(n-r)-rozmerná hrana je (n-r)-rozmernou podkockou kocky B^n.

Definícia 1.3.2 *Hrana n-rozmernej kocky B^n zodpovedajúca implikantu K (prostému implikantu) čiastočnej boolovskej funkcie f sa nazýva jej intervalom (maximálnym intervalom) a označuje N_K. Počet premenných, ktoré vystupujú v implikante K nazývame rádom intervalu a označujeme r(K).*

Interval r-tého rádu N_K indukuje (n-r)-rozmernú podkocku B^n. Tiež ak k_1 a k_2 sú elementárne konjunkcie, potom $N(k_1 \vee k_2) = N_{K_1} \cup N_{K_2}$.

Príklad 1.3.1 *Konjunkciám*

$$K_1(x_1, x_2, x_3) = \bar{x}_2 \wedge \bar{x}_3$$

$$K_2(x_1, x_2, x_3) = x_1 \wedge \bar{x}_2$$

$$K_3(x_1, x_2, x_3) = x_1$$

odpovedajú intervaly

$$N_{K_1} = \{(0,0,0), (1,0,0)\}$$

$$N_{K_2} = \{(1,0,0), (1,0,1)\}$$

$$N_{K_3} = \{(1,0,0),(1,0,1),(1,1,0),(1,1,1)\}$$

ktoré majú príslušné rády 2, 2 a 1. Tieto intervaly odpovedajú jednorozmernej hrane (N_{K_1}), jednorozmernej hrane (N_{K_2}) a dvojrozmernej hrane (N_{K_3}).

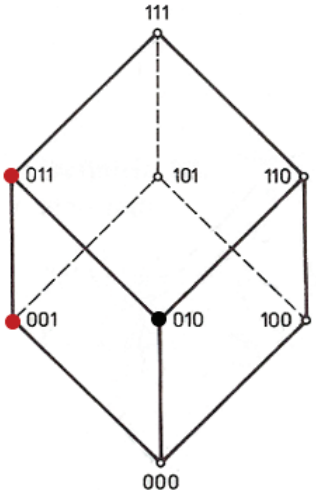
DNF čiastočnej boolovskej funkcie f odpovedá pokrytie N_f intervalmi $N_{K_1}, \ldots, N_{K_s}$ a každému pokrytiu množiny N_f intervalmi odpovedá DNF funkcie.

Z príkladu vyplýva, že existuje priama súvislosť medzi disjunktívnymi normálnymi formami čiastočnej boolovskej funkcie a vrcholovými pokrytiami množiny N_f intervalmi funkcie f. Z tohto pohľadu je štúdium čiastočných boolovských funkcií úzko spojené so štúdiom intervalových pokrytí náhodných indukovaných podgrafov v n-rozmernej kocke B^n.

Nech r_i označuje rád intervalu N_{K_i}, číslo r, kde $r = \sum_{i=1}^{s} r_i$, budeme nazývať rádom pokrytia.

Príklad 1.3.2 *Uvažujeme čiastočnú boolovskú funkciu $f(x_1, x_2, x_3)$ danú nasledujúcou Karnaughovou mapou*

Obr. 1.1: Projekcia 3-rozmernej kocky do roviny a Karnaughova mapa čiastočnej boolovskej funkcie f



x_1 \ x_2 x_3				
	1	0	-	-
x_1	1	1	1	1

Definičný obor funkcie f je množina $\{000, 010, 100, 101, 110, 111\}$. *Obor hodnôt funkcie f je množina* $\{0, 1\}$.

$$N_f = \{000, 100, 101, 110, 111\}$$

$$N_f^- = \{010\}$$

$$N_{\tilde{f}} = \{011, 001\}$$

Úplná DNF funkcie f:

$$U(f) = \bar{x_1}\bar{x_2}\bar{x_3} \vee x_1\bar{x_2}\bar{x_3} \vee x_1\bar{x_2}x_3 \vee x_1x_2\bar{x_3} \vee x_1x_2x_3$$

Skrátená DNF funkcie f:

$$C(f) = x_1 \vee \bar{x_2} \vee x_3$$

Ireducibilné[3] *pokrytie danej kocky je*

$$N_{\bar{x_2}} \cup N_{x_1}$$

kde

$$N_{\bar{x_2}} = \{(0,0,0),(0,0,1),(1,0,0)(1,0,1)\}$$

$$N_{x_1} = \{(1,0,0),(1,0,1),(1,1,0),(1,1,1)\}$$

Tieto intervaly majú príslušné rády 1, 1 a odpovedajú dvojrozmerným hranám. Rád pokrytia je 2.

Iredundantná DNF funkcie f:

$$T(f) = \bar{x_2} \vee x_1$$

Minimálna DNF funkcie f:

$$M(f) = \bar{x_2} \vee x_1$$

Najkratšia DNF funkcie f:

$$K(f) = \bar{x_2} \vee x_1$$

1.4 Skrátená DNF

Definícia 1.4.1 *Interval N_K sa nazýva maximálny, ak neexistuje interval N'_K taký, že*

1. $N_K \subseteq N'_K \subseteq (N_f \cup N_{\tilde{f}}) \wedge N_K \not\subseteq N_{\tilde{f}}$
2. *Rád intervalu N'_K je menší ako rád intervalu N_K.*

[3]Definované v Podkapitole 1.5.

Poznámka 1.4.1 *Konjunkcia K odpovedajúca maximálnemu intervalu N_K množiny N_f je prostým implikantom čiastočnej boolovskej funkcie f.*

Definícia 1.4.2 *DNF, ktorá je disjunkciou všetkých prostých implikantov funkcie f, sa nazýva skrátená DNF. Označujeme ju $C(f)$*

$$C(f) = K_1^0 \vee K_2^0 \vee \ldots \vee K_m^0$$

1.5 Iredundantnosť na základe geometrických znázornení

Definícia 1.5.1 *Pokrytie množiny N_f, pozostávajúce z maximálnych hrán sa nazýva ireducibilné, ak množina hrán, ktorá sa získa z pôvodnej vynechaním ľubovoľnej hrany, nebude pokrytím N_f.*

Definícia 1.5.2 *DNF, odpovedajúca ireducibilnému pokrytiu množiny N_f, sa nazýva iredundantná (v geometrickom zmysle).*

Pojem iredundantnej DNF vzhľadom na transformácie I a II a iredundantnej DNF v geometrickom zmysle sú ekvivalentné.

Medzi definovanými DNF - skrátenou, iredundantnou a minimálnou existujú nasledujúce vzťahy. Iredundantná DNF sa získa zo skrátenej vynechaním niektorých jej členov. Minimálna DNF (vzhľadom na index $l(D)$) je iredundantná. Medzi iredundantnými DNF sa nevyhnutne nachádza minimálna DNF (vzhľadom na index jednoduchosti).

Nech $M(f)$, $K(f)$ a $T(f)$ označujú množinu minimálnych, najkratších a iredundantných DNF čiastočnej boolovskej funkcie f. Nech $l^C(f) = l(C(f))$ je dĺžka skrátenej DNF, $L^M(f)$ je zložitosť minimálnej DNF, $l^K(f)$ je dĺžka

najkratšej DNF a $l^T(f)$ je najväčšia z dĺžok iredundantných DNF, potom označíme:

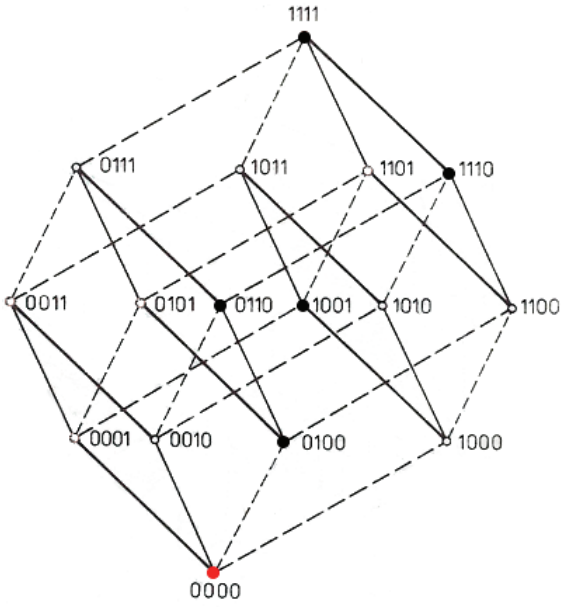
$$l_Q^C(n) = \max_{f \in Q} l^C(f), \qquad l_Q^T(n) = \max_{f \in Q} l^T(f)$$

$$l_Q^K(n) = \max_{f \in Q} l^K(f), \qquad l_Q^M(n) = \max_{f \in Q} L^M(f)$$

označujeme v uvedenom poradí maximálne hodnoty dĺžok skrátených, iredundantných, najkratších a maximálnu hodnotu zložitosti minimálnych DNF na množine $Q \subseteq P^n$.

Príklad 1.5.1 *Uvažujeme čiastočnú boolovskú funkciu* $f(x_1, x_2, x_3, x_4)$ *danú nasledujúcou Karnaughovou mapou*

Obr. 1.2: Projekcia 4-rozmernej kocky do roviny a Karnaughova mapa čiastočnej boolovskej funkcie f



x_1 \ x_3 x_4			
-	1	1	1
1	1	1	0
1	0	0	1
0	0	1	1

Definičný obor funkcie f je množina $\{1111, 1110, 1101, 1100, 1011, 1010, 1001, 1000, 0111, 0110, 0101, 0100, 0011, 0010, 0001\}$. *Obor hodnôt funkcie f je množina* $\{0, 1\}$.

$$N_f = \{1101, 1100, 1011, 1010, 1000, 0111, 0101, 0011, 0010, 0001\}$$

$$N_f^- = \{1111, 1110, 1001, 0110, 0100\}$$

$$N_{\tilde{f}} = \{0000\}$$

Maximálnym hranám $N_1, \ldots, N_7$ *odpovedajú prosté implikanty*

$$K_1 = \bar{x}_2\bar{x}_4, \quad K_2 = \bar{x}_1\bar{x}_2, \quad K_3 = \bar{x}_1x_4, \quad K_4 = \bar{x}_2x_3$$

$$K_5 = x_2\bar{x}_3x_4, \;\; K_6 = x_1x_2\bar{x}_3, \;\; K_7 = x_1\bar{x}_3\bar{x}_4$$

Skrátená DNF:

$$C(f) = \bar{x}_2\bar{x}_4 \vee \bar{x}_1\bar{x}_2 \vee \bar{x}_1x_4 \vee \bar{x}_2x_3 \vee x_2\bar{x}_3x_4 \vee x_1x_2\bar{x}_3 \vee x_1\bar{x}_3\bar{x}_4$$

Dve hrany - N_3 a N_4 patria do ľubovoľného pokrytia, pretože iba ony pokrývajú vrcholy (0111) a (1011). Na pokrytie vrcholu (1000) treba vziať hranu N_1 alebo hranu N_7 a na pokrytie vrcholov (1100), (1101) treba vziať hranu N_6 alebo dvojicu hrán N_5 a N_7.

Iredundantné DNF, ktoré dostaneme:

$$T_1 = \bar{x}_2\bar{x}_4 \vee \bar{x}_1x_4 \vee \bar{x}_2x_3 \vee x_2\bar{x}_3x_4 \vee x_1\bar{x}_3\bar{x}_4$$

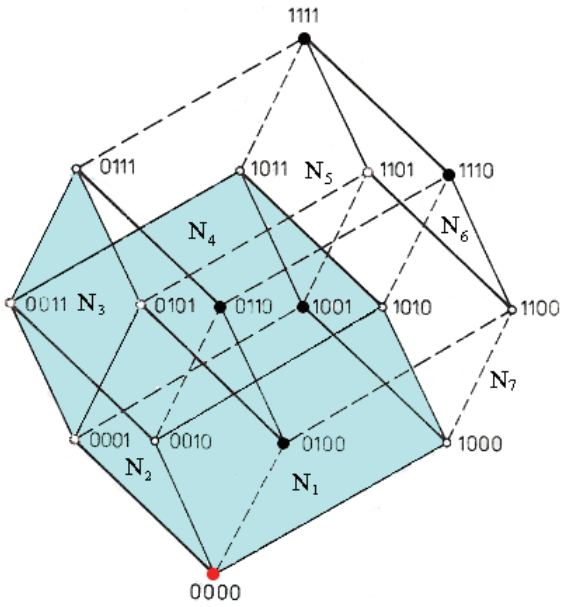
$$T_2 = \bar{x}_2\bar{x}_4 \vee \bar{x}_1x_4 \vee \bar{x}_2x_3 \vee x_1x_2\bar{x}_3$$

$$T_3 = \bar{x}_1x_4 \vee \bar{x}_2x_3 \vee x_2\bar{x}_3x_4 \vee x_1\bar{x}_3\bar{x}_4$$

$$T_4 = \bar{x}_1x_4 \vee \bar{x}_2x_3 \vee x_1x_2\bar{x}_3 \vee x_1\bar{x}_3\bar{x}_4$$

$$l(T_1) = 12, \;\; l(T_2) = 9, \;\; l(T_3) = l(T_4) = 10$$

Obr. 1.3: Množina N_f a maximálne hrany



1.6 Pravdepodobnosť

Definícia 1.6.1 *Nech Ω je neprázdna množina a nech $S \subseteq 2^{\Omega}$. Usporiadanú dvojicu (Ω, S) nazývame Ω-algebra, alebo Ω-algebra udalostí, ak platí:*

1. $\Omega \in S$
2. $A \in S \Rightarrow \Omega \setminus A \in S$
3. *Ak $(A_i)_{i \in I}$ je spočítateľná postupnosť prvkov systému S, tak $\bigcup_{i \in I} A_i \in S$*

Definícia 1.6.2 *Nech (Ω, S) je Ω-algebra. Prvky množiny Ω nazveme elementárne výsledky a podmnožiny Ω patriace do systému S nazveme udalosti.*

Definícia 1.6.3 *Pravdepodobnostná miera na Ω-algebre udalostí (Ω, S) je zobrazenie $P : S \to \mathbb{R}$ spĺňajúce nasledovné podmienky:*

1. *Pre všetky* $A \in S$ *platí* $0 \leq \mathrm{P(A)} \leq 1$

2. $\mathrm{P}(\Omega) = 1$, $\mathrm{P}(\emptyset) = 0$

3. *Ak* $(A_i)_{i \in I}$ *je postupnosť disjunktných udalostí, tak* $\mathrm{P}(\bigcup_{\mathrm{i} \in \mathrm{I}} \mathrm{A_i}) = \sum_{\mathrm{i} \in \mathrm{I}} \mathrm{P(A_i)}$

Definícia 1.6.4 *Nech* (Ω, S) *je* Ω*-algebra udalostí a nech* P *je pravdepodobnostná miera udalostí na* (Ω, S)*. Potom trojicu* (Ω, S, P) *nazveme pravdepodobnostný priestor.*

Definícia 1.6.5 *Nech* (Ω, S, P) *je pravdepodobnostný priestor. Budeme hovoriť, že funkcia* $\mathrm{X} : \Omega \to \mathbb{R}$ *je náhodná premenná, ak pre každé* $x \in \mathbb{R}$ *platí*

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

Definícia 1.6.6 *Náhodnú premennú* X *na pravdepodobnostnom priestore* (Ω, S, P) *nazývame diskrétna, ak jej obor hodnôt* $\mathrm{X}(\Omega) \subset \mathbb{R}$ *je spočítateľná množina.*

Definícia 1.6.7 *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*

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

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

Veta 1.6.1 *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*

$$\mathrm{E(aX + bY) = aE(X) + bE(Y)}$$

Špeciálne, $\mathrm{E(aX) = aE(X)}$ *a* $E(X + Y) = E(X) + E(Y)$*.*

Definícia 1.6.8 *Nech diskrétna náhodná premenná* X *má konečnú strednú hodnotu. Disperziou náhodnej premennej* X *nazveme hodnotu*

$$\mathrm{D(X)} = \mathrm{E}((\mathrm{X} - \mathrm{E(X)})^2)$$

Veta 1.6.2 *Nech X je diskrétna náhodná premenná s konečnou strednou hodnotou aj disperziou. Potom platí:*

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

Veta 1.6.3 *(Markovova nerovnosť). Nech X je nezáporná náhodná premenná. Potom pre každé $a > 0$ platí*

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

Veta 1.6.4 *(Čebyševova nerovnosť). Nech X je diskrétna náhodná premenná s konečnou disperziou (t.j. aj s konečnou strednou hodnotou). Potom pre každé $b > 0$ platí*

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

V tejto diplomovej práci budeme uvažovať všetky náhodné premenné celočíselné a nezáporné.

1.7 Asymptotické ohraničenia

Nech S je istá vlastnosť a f je náhodná čiastočná boolovská funkcia. Ak $\lim_{n\to\infty} P[\text{f má vlastnosť S}] = 1$ budeme hovoriť, že funckia f má vlastnosť S alebo že funkcia f spĺňa vlastnosť S takmer určite. Budeme používať O-notáciu a o-notáciu. Symbol $o(a_n)$ označuje výraz, ktorý ide k 0, keď je delený a_n. Symbol $O(a_n)$ označuje výraz, ktorý ostáva ohraničený, keď je

delený a_n. Postupnosti (a_n) a (b_n) sú asymptoticky ekvivalentné, ak $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$. Asymptotickú ekvivalenciu označujeme $a_n \sim b_n$. Symbol $\log x$ označuje algoritmus pri základe 2.

Kapitola 2

Odhady niektorých parametrov čiastočných boolovských funkcií

V tejto kapitole popíšeme vzťah medzi parametrami DNF všade definovaných a čiastočných boolovských funkcií. Ďalej definujeme rozptyl dĺžok a zložitosti iredundantných DNF a uvedieme výsledky L. V. Vasielieva, ktorý asymptoticky odhadol tento rozptyl pre úplné boolovské funkcie. My tieto dva parametre asymptoticky odhadneme pre čiastočné boolovské funkcie. Tiež uvedieme výsledky Lin Sin-Ljana, ktorý porovnal zložitosť najkratších a minimálnych DNF pre úplné boolovské funkcie.

2.1 Vzťah medzi parametrami DNF všade definovaných a čiastočných boolovských funkcií

Nech $\chi(f)$, $\tilde{\chi}(f)$ je niektorý číselný parameter definovaný na množine P^n, $\tilde{P}^n$. Označíme $\chi_{\tilde{P}}(n) = \max\limits_{f \in \tilde{P}^n} \tilde{\chi}(f)$ a $\chi_P(n) = \max\limits_{f \in P^n} \chi(f)$.

Poznámka 2.1.1 *Pretože $\tilde{P}^n \subset P^n$, potom zrejme $\chi_{\tilde{P}}(n) \leq \chi_P(n)$. (1)*

Budeme sa snažiť ukázať, že pre základné parametre charakterizujúce zložitosť DNF nerovnosť (1) sa mení na rovnosť. Na druhej strane musíme zistiť, kedy pri niektorých parametroch platí ostrá nerovnosť.

Nech $f \in P^n$ a D je niektorá DNF, ktorá realizuje čiastočnú boolovskú funkciu f. F_D budeme označovať všade definovanú boolovskú funkciu realizovanú disjunktívnou normálnou formou D.

Lema 2.1.0.1 *Nech $f \in P^n$ a D je jej skrátená DNF. Potom každý prostý implikant funkcie f je aj prostý implikant pre všade definovanú funkciu $F_D = \varphi$.*

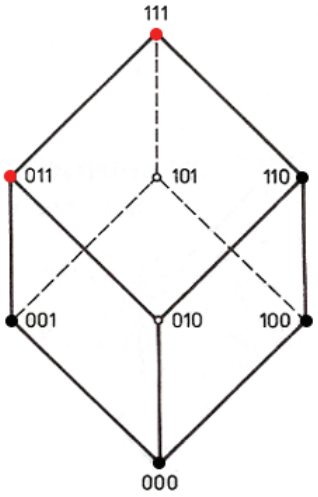
Dôkaz 2.1.1 *Platí, že $N_f \subseteq N_\varphi$, $N_f^- \subseteq N_\varphi^-$. Nech K je prostým implikantom funkcie f. Z definície funkcie φ vyplýva, že $N_\varphi \cap N_K \neq \emptyset$, $N_\varphi^- \cap N_K = \emptyset$, t.j. K je implikantom funkcie φ. Ukážeme, že nie je možné z K vynechať žiadne činitele. Nech K' je konjunkcia, ktorú dostaneme z K vynechaním niektorého činiteľa. Dostávame $N_{K'} \cap N_f^- \neq \emptyset$, pretože K je prostý implikant funkcie f. Ale $N_f^- \subseteq N_\varphi^-$. Z toho vyplýva, že K' nie je implikantom funkcie φ. Lema je týmto dokázaná.*

Z predchádzajúcej lemy vyplýva, že $l^C(f) \leq l^C(F_{C(f)})$.

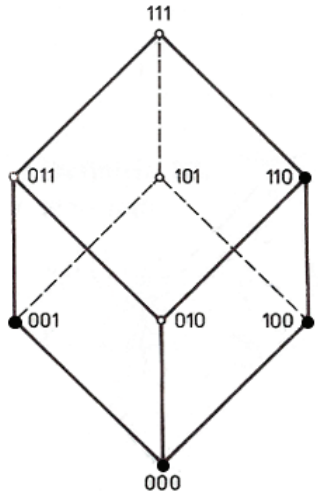
Nasledujúci príklad ukazuje, že táto nerovnosť môže byť ostrá pre niektoré funkcie.

Príklad 2.1.1 *Uvažujeme funkciu $f(x_1, x_2, x_3) \in P^3$ pre ktorú $N_f = (0,1,0), (1,0,1)$, $N_{\tilde{f}} = (0,1,1), (1,1,1)$ a funkciu $\varphi(x_1, x_2, x_3)$ pre ktorú $N_\varphi = (0,1,0), (1,0,1), (1,1,1), (0,1,1)$.*

Obr. 2.1: Projekcia 3-rozmernej kocky do roviny, funkcie f a φ a príslušné Karnaughove mapy



x_1 \ x_2 x_3	00	01	11	10
0	0	1	-	0
1	0	0	-	1



x_1 \ x_2 x_3	00	01	11	10
0	0	1	1	0
1	0	0	1	1

Potom $D = C(f) = \bar{x}_1 x_2 \vee x_1 x_3$ a súčasne skrátená DNF funkcie $F_D = \varphi$ má tvar $C(F_D) = \bar{x}_1 x_2 \vee x_1 x_3 \vee x_2 x_3$, teda $l^C(f) < l^C(F_{C(f)})$.

Lema 2.1.0.2 *Nech D je iredundantná DNF čiastočnej boolovskej funkcie $f \in P^n$. Potom D je iredundantná DNF pre funkciu F_D.*

Dôkaz 2.1.2 *Z predchádzajúcej lemy vyplýva, že každá konjunkcia DNF D je prostým implikantom funkcie F_D. Z toho vyplýva, že zo žiadnej konjunkcie nemôžeme vynechať činiteľa. Ukážeme, že nemožno vynechať konjunkciu. Po vynechaní ľubovoľnej konjunkcie z D dostávame DNF D', ktorá nerealizuje funkciu f. Táto nová DNF D' nemôže realizovať ani funkciu F_D, pretože $N_{D'} \subseteq N_f \subseteq N_{F_D}$. Teda z D nemôžeme vynechať žiadnu konjunkciu.*

Lema 2.1.0.3 *Ak D je minimálna (najkratšia) DNF funkcie $f \in P^n$, potom D je minimálna (najkratšia) DNF funkcie F_D.*

Dôkaz 2.1.3 *Sporom. Nech D_1 je minimálna DNF funkcie F_D, pričom $l(D_1) < l(D)$. Uvažujme DNF D_2 zloženú zo všetkých konjunkcií K, vystupujúcich v D_1 a takých, že $N_K \cap N_f \neq \emptyset$. Samozrejme, že $N_f \subseteq N_{D_2}$. To znamená, že D_2 realizuje a má menšiu zložitosť ako D. Dostali sme sa do sporu, čo dokazuje túto lemu.*

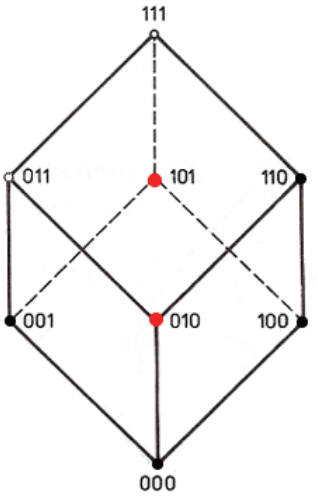
Vzniká otázka, či platí obrátené tvrdenie k tejto leme.

Tvrdenie 2.1.1 *Nech D je minimálna (najkratšia) DNF všade definovanej funkcie φ. Nech f je ľubovoľná čiastočná boolovská funkcia, pre ktorú D je iredundantná DNF. Potom D je minimálna (najkratšia) DNF funkcie f.*

Toto tvrdenie vo všeobecnosti neplatí, ako ukazuje nasledujúci príklad.

Príklad 2.1.2 *Nech φ je všade definovaná funkcia, ktorá realizuje DNF $D = \bar{x}_1 x_2 \vee x_1 x_3$. DNF D je minimálnou a najkratšou pre funkciu φ. Nech f je čiastočná funkcia taká, že $N_f = \{(1,1,1),(0,1,1)\}$, $N_{\tilde{f}} = \{(1,0,1),(0,1,0)\}$.*

Obr. 2.2: Projekcia 3-rozmernej kocky do roviny, funkcia f



x_1 \ x_2 x_3				
	0	-	1	0
	0	0	1	-

DNF D je iredundantná pre funkciu f avšak najkratšia a minimálna DNF pre funkciu f je $D_1 = x_2x_3$.

Veta 2.1.1 *Nech* $\chi(f)$ *je ľubovoľný z parametrov* $l^C(f), l^T(F), l^K(f), L^M(f)$. *Potom* $\chi_P(n) = \chi_{\tilde{P}}(n)$.

Dôkaz 2.1.4 *Z predchádzjúcich liem vyplýva, že pre ľubovoľný parameter* χ *uvedený vo vete a ľubovoľnej funkcii* $f \in P^n$ *existuje funkcia* $\varphi \in P^n$ *taká, že* $\chi(f) \leq \chi(\varphi)$. *Odtiaľ a z poznámky 2.1.1 vyplýva tvrdenie vety.*

2.2 Rozptyl dĺžok a zložitosti iredundantných DNF

V práci [7] sa uvažovali vzťahy medzi dĺžkami a zložitosťami iredundantných DNF realizujúcich jednu a tú istú všade definovanú funkciu algebry logiky.

Definícia 2.2.1 *Nech f je ľubovoľná čiastočná boolovská funkcia. Veličina*

$$Y(f) = \max_{D_1, D_2 \in T(f)} \frac{L(D_1)}{L(D_2)}$$

sa nazýva rozptyl dĺžok.

Definícia 2.2.2 *Nech f je ľubovoľná čiastočná boolovská funkcia. Veličina*

$$R(f) = \max_{D_1, D_2 \in T(f)} \frac{l(D_1)}{l(D_2)}$$

sa nazýva rozptyl zložitosti iredundantných DNF funkcie f.

Nech $Q \subseteq P^n$. Potom $Y_Q(n) = \max_{f \in Q} Y(f)$, $R_Q(n) = \max_{f \in Q} R(f)$. V práci [7] Y.L.Vasiliev dokázal, že

$$2^{n - \sqrt[3]{n}} \leq Y_{\tilde{P}}(n) \leq 2^{n - \log n}$$

$$2^{n - \sqrt[3]{n}} \leq R_{\tilde{P}}(n) \leq 2^n$$

Ukazuje sa, že pre množiny nie všade definovaných funkcií môžeme získať asymptotické odhady.

Veta 2.2.1 $Y_P(n) \sim 2^{n-1}$, $R_P(n) \sim n * 2^{n-1}$

Dôkaz 2.2.1 *Odhad zhora pre $Y_P(n)$. Nech funkcia $f \in P^n$ taká, že $Y(f) = Y_P(n)$ a nech D_1, D_2 sú iredundantné DNF funkcie f také, že $Y_P(n) = Y(f) = L(D_1)/L(D_2)$.*

Môžu nastať dva prípady:

1. *$L(D_2) = 1$. Potom existuje hrana g kocky B^n, obsahujúca všetky vrcholy z N_f. Hodnosť tejto hrany je väčšia ako 0, pretože v opačnom prípade funkcia má jedinú iredundantnú DNF a z toho vyplýva, že $Y(f) = 1$. Nech hrana g má hodnosť $r \geq 1$, potom*

$$Y_P(n) = Y(f) = L(D_1) \leq |N_f| \leq 2^{n-r} \leq 2^{n-1}$$

2. *$L(D_2) \geq 2$. Potom $Y_P(n) = Y(f) \leq |N_f|/L(D_2) \leq 2^{n-1}$. Odtiaľ vyplýva, že $Y_P(n) \leq 2^{n-1}$.*

Odhad zhora pre $R_P(n)$. Nech funkcia $f \in P^n$ je taká, že $R(f) = R_P(n)$ a nech D_1, D_2 sú iredundantné DNF funkcie f také, že $R(f) = l(D_1)/l(D_2)$. $N_f^- \neq \emptyset$ a teda ľubovoľný implikant funkcie f má hodnosť väčšiu ako 0. Potom

$$R_P(n) = R(f) = \frac{l(D_1)}{l(D_2)} \leq \frac{L(D_1) \max\limits_{K_1 \in D_1} r(K_1)}{L(D_2) \min\limits_{K_2 \in D_2} r(K_2)} \leq Y_P(n)n \leq n2^{n-1}$$

kde r(K) je hodnosť konjunkcie K. A teda

$$R_P(n) \leq n2^{n-1}$$

Dolné odhady. Uvažujme ľubovoľnú všade definovanú funkciu $\varphi(x_1, x_2, \ldots, x_{n-1})$. Definujeme čiastočnú funkciu $f_\varphi(x_1, x_2, \ldots, x_n)$ nasledujúcim spôsobom:

$$f_\varphi(\alpha_1, \ldots, \alpha_{n-1}, 0) = 0, \;\; ak \;\; \varphi(\alpha_1, \ldots, \alpha_{n-1}) = 0$$

$$f_\varphi(\alpha_1, \ldots, \alpha_{n-1}, 1) = 1, \;\; ak \;\; \varphi(\alpha_1, \ldots, \alpha_{n-1}) = 1$$

Na ostatných súboroch $(\alpha_1, \ldots, \alpha_n)$ funkcia f_φ nie je definovaná. Funkcia má nasledujúce vlastnosti:

1. *DNF x_n realizuje f_φ a je minimálnou a najkratšou.*

2. *Každú iredundantnú DNF funkcie φ je iredundantná DNF funkcie f_φ*

V práci V.V.Glogoljeva [6] je zostrojená pre ľubovoľné (dostatočne veľké) n funkcia $\varphi(x_1, x_2, \ldots, x_{n-1})$, pre ktorú existuje iredundantná DNF D taká, že $L(D) \sim n2^{n-1}$, $l(D) \sim n2^{n-1}$. Odtiaľ vyplýva, že pre zodpovedajúcu funkciu f_φ platí:

$$Y(f_\varphi) \geq 2^{n-1}(1 - \delta'_n), \;\; R(f_\varphi) \geq n2^{n-1}(1 - \delta''_n), \;\; \delta'_n, \delta''_n \to 0, n \to \infty$$

2.3 Porovnanie zložitosti najkratších a minimálnych DNF

V práci Lin Sin-Ljana [5] sa študovali vzťahy medzi zložitosťami rôznych najkratších DNF jednej a tej istej všade definovanej funkcii. Skúma sa taktiež otázka vzťahu zložitosti najkratších a minimálnych DNF. Nech

$$U(f) = \max_{K_1, K_2 \in K(f)} \frac{l(K_1)}{l(K_2)}$$

kde maximum uvažujeme pre všetky dvojice najkratších DNF funkcie f. Nech

$$V(f) = min \frac{l(K)}{l(f_m)}, \;\; kde f_m \in M(f)$$

V práci [5] je ukázané, že

$$\max_{f \in \tilde{P}^n} V(f) \sim \max_{f \in \tilde{P}^n} U(f) \sim \frac{n}{2}$$

Kapitola 3

Odhady niektorých parametrov pre skoro všetky čiastočné boolovské funkcie

V tejto kapitole najprv popíšeme pravdepodobnostný priestor, v ktorom budeme pracovať. Ďalej sa budeme venovať asymptotickému odhadu počtu k-rozmerných intervalov čiastočnej náhodnej boolovskej funkcie a dĺžke úplnej DNF funkcie. Následne asymptoticky odhadneme počet k-rozmerných maximálnych intervalov a dĺžku skrátenej DNF.

3.1 Pravdepodobnostný priestor

Náhodné čiastočné boolovské funkcie tvoria diskrétny pravdepodobnostný priestor $(P^n, 2^{P^n}, P)$, kde P^n je množina n-árnych náhodných čiastočných boolovských funkcií a P je pravdepodobnosť, ktorej definícia je motivovaná nasledujúcimi úvahami.

Náhodný výber funkcie $f \in P^n$ si môžeme predstaviť ako náhodné rozdelenie prvkov B^n do troch disjunktných množín - množiny N_f, N_f^- a $N_{\tilde{f}}$,

pričom pre každú n-ticu $\alpha \in B^n$ platí, že sa zobrazí na hodnotu 1 s pravdepodobnosťou p_1, na hodnotu 0 s pravdepodobnosťou p_2 a nebude definovaná s pravdepodobnosťou p_3, pričom $p_1 + p_2 + p_3 = 1$. Pravdepodobnosť výberu funkcie $f \in P^n$ je rovné

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

Teda pre ľubovoľnú podmnožinu čiastočných boolovských funkcií $A \subseteq P^n$ platí

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

Veta 3.1.1 *Nech U, V a W sú po dvoch disjunktné podmnožiny $B^n = \{0,1\}^n$ a $F = \{f \in P^n; U \subseteq N_f; V \subseteq N_f^-; W \subseteq N_{\tilde{f}}\}$. Potom*

$$P[F] = P[\{f \in P^n;\ U \subseteq N_f \wedge V \subseteq N_f^- \wedge W \subseteq N_{\tilde{f}}\}] = p_1^{|U|} \cdot p_2^{|V|} \cdot p_3^{|W|}$$

Dôkaz 3.1.1 *Označme si $a = 2^n - |U \cup V \cup W|$.*

$$P[F] = \sum_{k=0}^{a} \binom{a}{k} \cdot p_1^{|U|+k} \cdot \sum_{j=0}^{a-k} \binom{a-k}{j} \cdot p_2^{|V|+j} \cdot p_3^{|W|+a-k-j} =$$

$$= \sum_{k=0}^{a} \binom{a}{k} \cdot p_1^{|U|+k} \cdot p_2^{|V|} \cdot p_3^{|W|} \cdot \sum_{j=0}^{a-k} \binom{a-k}{j} p_2^j \cdot p_3^{a-k-j} =$$

$$= \sum_{k=0}^{a} \binom{a}{k} \cdot p_1^{|U|+k} \cdot p_2^{|V|} \cdot p_3^{|W|} \cdot (p_2 + p_3)^{a-k} =$$

$$= p_1^{|U|} \cdot p_2^{|V|} \cdot p_3^{|W|} \cdot \sum_{k=0}^{a} \binom{a}{k} p_1^k \cdot (p_2 + p_3)^{a-k} =$$

$$= p_1^{|U|} \cdot p_2^{|V|} \cdot p_3^{|W|} \cdot (p_1 + p_2 + p_3)^a = p_1^{|U|} \cdot p_2^{|V|} \cdot p_3^{|W|}$$

Veta 3.1.2 *Nech K je k-rozmerná podkocka B^n, potom*

$$P[K \subseteq (N_f \cup N_{\tilde{f}}) \wedge K \not\subseteq N_{\tilde{f}}] = P[K \subseteq (N_f \cup N_{\tilde{f}})] - P[K \subseteq N_{\tilde{f}}] = (p_1 + p_3)^{2^k} - p_3^{2^k}$$

3.2 Intervaly

V nasledujúcej podkapitole určíme strednú hodnotu a disperziu počtu k-rozmerných intervalov náhodnej funkcie f z P^n, vypočítame dolný a horný odhad náhodnej premennej označujúcej tento počet. Ďalej asymptoticky odhadneme počet k-rozmerných intervalov čiastočnej boolovskej funkcie, čo nám umožní asymptoticky odhadnúť počet bodov funkcie, ktoré sa zobrazia na hodnotu 1, teda dĺžku úplnej DNF čiastočnej boolovskej funkcie.

Nech $i_{n,k}$ označuje náhodnú premennú na P^n takú, že $i_{n,k}(f)$ je rovné počtu k-rozmerných intervalov funkcie f z P^n.

Veta 3.2.1 $Ei_{n,k} = \binom{n}{k} \cdot 2^{n-k} \cdot ((p_1 + p_3)^{2^k} - p_3^{2^k})$

Dôkaz 3.2.1 *Majme ľubovoľnú k-rozmernú podkocku K kocky B^n. Nech η_K je náhodná premenná (alebo aj indikátor), definovaná nasledovne:*

$$\eta_K = \begin{cases} 1 & ak\ K \subseteq (N_f \cup N_{\tilde{f}}) \wedge K \not\subseteq N_{\tilde{f}} \\ 0 & inak \end{cases}$$

Potom pre náhodnú premennú $i_{n,k}$ platí :

$$i_{n,k} = \sum_K \eta_K$$

kde K symbolizuje všetky k-rozmerné podkocky B^n.
Z Vety 3.1.2 vyplýva:

$$E\eta_K = P[K \subseteq (N_f \cup N_{\tilde{f}}) \wedge K \not\subseteq N_{\tilde{f}}] = P[\eta_K = 1] = (p_1 + p_3)^{2^k} - p_3^{2^k}$$

Keďže v B^n existuje $\binom{n}{k} \cdot 2^{n-k}$ k-rozmerných podkociek, tak musí platiť:

$$Ei_{n,k} = \sum_K E\eta_K = \binom{n}{k} \cdot 2^{n-k} \cdot ((p_1 + p_3)^{2^k} - p_3^{2^k})$$

Veta 3.2.2

$$D_{i_{n,k}} \leq 2^{n-k} \cdot \binom{n}{k}^2 \cdot \left((p_1 + p_3)^{2^k} - p_3^{2^k} \right)$$

Dôkaz 3.2.2 *Dôkaz sa skladá z výpočtu $E(i_{n,k})^2$ a $(Ei_{n,k})^2$. Náhodnú premennú $i_{n,k}$ môžme vyjadriť ako sumu indikátorov η_K a teda*

$$(i_{n,k})^2 = \left(\sum_K \eta_K \right)^2 = \sum_{(K,L)} \eta_K \cdot \eta_L$$

kde (K, L) sú všetky usporiadané dvojice k-rozmerných podkociek B^n.
Môžu nastať 2 možnosti

1. $K \cap L \neq \emptyset$
   Nech prienikom je j-rozmerná podkocka, $0 \leq j \leq k$, $|K \cup L| = 2^{k+1} - 2^j$.
   Keďže obe kocky musia obsahovať aspoň jednu jednotku, pre výpočet pravdepodobnosti využijeme princíp inklúzie a exklúzie.

   $$P[\eta_K \cdot \eta_L = 1] = (p_1+p_3)^{2^{k+1}-2^j} - 2 \cdot p_3^{2^k} \cdot (p_1+p_3)^{2^k-2^j} + p_3^{2^{k+1}-2^j} = E(\eta_K \cdot \eta_L)$$

2. $K \cap L = \emptyset$, $|K \cup L| = 2^{k+1}$

   $$P[\eta_K \cdot \eta_L = 1] = (p_1 + p_3)^{2^{k+1}} - 2 \cdot p_3^{2^k} \cdot (p_1 + p_3)^{2^k} + p_3^{2^{k+1}} = E(\eta_K \cdot \eta_L)$$

Počet dvojíc (K, L) s prienikom dimenzie j je

$$2^{n-j} \cdot \binom{n}{j} \cdot \binom{n-j}{k-j} \cdot \binom{n-k}{k-j}$$

a počet takých dvojíc (K, L), ktoré nemajú spoločný prienik je:

$$\left(\binom{n}{k} \cdot 2^{n-k} \right)^2 - \sum_{j=0}^{k} 2^{n-j} \cdot \binom{n}{j} \cdot \binom{n-j}{k-j} \cdot \binom{n-k}{k-j}$$

Z predchádzajúcich výsledkov dostávame

$$E(i_{n,k})^2 =$$

$$= \sum_{j=0}^{k} 2^{n-j} \binom{n}{j} \binom{n-j}{k-j} \binom{n-k}{k-j} ((p_1+p_3)^{2^{k+1}-2^j} - 2 \cdot p_3^{2^k} \cdot (p_1+p_3)^{2^k-2^j} + p_3^{2^{k+1}-2^j}) +$$

$$+ (((\binom{n}{k} \cdot 2^{n-k})^2 - \sum_{j=0}^{k} 2^{n-j} \binom{n}{j} \binom{n-j}{k-j} \binom{n-k}{k-j}))((p_1+p_3)^{2^{k+1}} - 2p_3^{2^k}(p_1+p_3)^{2^k} + p_3^{2^{k+1}}) =$$

$$= 2^n \cdot \binom{n}{k} \cdot \sum_{j=0}^{k} \binom{k}{j} \binom{n-k}{k-j} 2^{-j} ((p_1+p_3)^{2^{k+1}-2^j} - (p_1+p_3)^{2^{k+1}} - 2p_3^{2^k}(p_1+p_3)^{2^k-2^j} +$$

$$2p_3^{2^k}(p_1+p_3)^{2^k} + p_3^{2^{k+1}-2^j} - p_3^{2^{k+1}}) + (\binom{n}{k} \cdot 2^{n-k} \cdot ((p_1+p_3)^{2^k} - p_3^{2^k}))^2$$

$$E^2(i_{n,k}) = \left(\binom{n}{k} \cdot 2^{n-k} \cdot \left((p_1+p_3)^{2^k} - p_3^{2^k} \right) \right)^2$$

Z predošlých vzťahov teraz môžeme vyjadriť a ohraničiť disperziu.

$$D_{i_{n,k}} = E(i_{n,k})^2 - E^2(i_{n,k})$$

$$D_{i_{n,k}} = 2^n \cdot \binom{n}{k} \cdot \sum_{j=0}^{k} \binom{k}{j} \cdot \binom{n-k}{k-j} \cdot 2^{-j} \cdot ((p_1+p_3)^{2^{k+1}-2^j} - (p_1+p_3)^{2^{k+1}} - 2p_3^{2^k}(p_1+$$

$$p_3)^{2^k-2^j} + 2p_3^{2^k}(p_1+p_3)^{2^k} + p_3^{2^{k+1}-2^j} - p_3^{2^{k+1}})$$

Aby sme vedeli zhora daný výraz ohraničiť, potrebujeme ukázať neklesajúcosť výrazu

$$2^{-j} \cdot ((p_1+p_3)^{2^{k+1}-2^j} - (p_1+p_3)^{2^{k+1}} - 2p_3^{2^k}(p_1+p_3)^{2^k-2^j} + 2p_3^{2^k}(p_1+p_3)^{2^k} +$$

$$p_3^{2^{k+1}-2^j} - p_3^{2^{k+1}}) = a_j$$

Chceme ukázať, že $a_j \leq a_{j+1}$.

$$2^{-j} \cdot ((p_1+p_3)^{2^{k+1}-2^j} - (p_1+p_3)^{2^{k+1}} - 2p_3^{2^k}(p_1+p_3)^{2^k-2^j} + 2p_3^{2^k}(p_1+p_3)^{2^k} +$$

$$p_3^{2^{k+1}-2^j} - p_3^{2^{k+1}}) \leq \frac{2^{-j}}{2} \cdot ((p_1+p_3)^{2^{k+1}-2^{j+1}} - (p_1+p_3)^{2^{k+1}} - 2 \cdot p_3^{2^k} \cdot (p_1+$$

$$p_3)^{2^k-2^{j+1}} + 2 \cdot p_3^{2^k} \cdot (p_1+p_3)^{2^k} + p_3^{2^{k+1}-2^{j+1}} - p_3^{2^{k+1}})$$

$0 \leq \frac{1}{2}((p_1+p_3)^{2^{k+1}}(((p_1+p_3)^{-2^{j+1}}-2(p_1+p_3)^{-2^j}+1)-2p_3^{2^k}(p_1+p_3)^{2^k}((p_1+p_3)^{-2^{j+1}}-2(p_1+p_3)^{-2^j}+1)+p_3^{2^{k+1}}(p_3^{-2^{j+1}}-2p_3^{-2^j}+1))=a$

Teraz ukážeme, že ak posledný výraz $(p_3^{-2^{j+1}}-2p_3^{-2^j}+1)$ *nahradíme výrazom*
$((p_3+p_1)^{-2^{j+1}}-2(p_3+p_1)^{-2^j}+1)$, *celý výraz zmenšíme*

$$p_3^{-2^{j+1}}-2p_3^{-2^j}+1 \geq (p_3+p_1)^{-2^{j+1}}-2(p_3+p_1)^{-2^j}+1$$

$$(p_3^{-2^j}-1)^2 \geq ((p_3+p_1)^{-2^j}-1)^2$$

Keďže $p_3, p_1+p_3 \in (0,1)$, *potom* $p_3^{-2^j}, (p_1+p_3)^{-2^j} \in (1,\infty)$. *Vieme, že* $p_1+p_3 \geq p_3$, *čiže* $(p_1+p_3)^{-2^j} \leq p_3^{-2^j}$ *Z toho vyplýva, že zamenením* p_3 *za* p_1+p_3 *v poslednej zátvorke daný výraz zmenšíme.*

$a \geq \frac{1}{2}((p_1+p_3)^{2^{k+1}}((p_1+p_3)^{-2^{j+1}}-2(p_1+p_3)^{-2^j}+1)-2p_3^{2^k}(p_1+p_3)^{2^k}((p_1+p_3)^{-2^{j+1}}-2(p_1+p_3)^{-2^j}+1)+p_3^{2^{k+1}}((p_1+p_3)^{-2^{j+1}}-2(p_3+p_1)^{-2^j}+1)) \geq 0$

$$a \geq \frac{1}{2}((p_1+p_3)^{2^k}-p_3^{2^k})^2((p_1+p_3)^{-2^j}-1)^2 \geq 0$$

Z nezápornosti druhých mocnín je zrejmé, že výraz $a \geq 0$, *čím sme ukázali neklesajúcosť* a_j.

Teraz môžeme $D_{i_{n,k}}$ *zhora ohraničiť*

$$D_{i_{n,k}} = 2^n \cdot \binom{n}{k} \cdot \sum_{j=0}^{k} \binom{k}{j} \cdot \binom{n-k}{k-j} a_j \leq 2^n \cdot \binom{n}{k} \cdot a_k \cdot \sum_{j=0}^{k} \binom{k}{j} \cdot \binom{n-k}{k-j} = 2^n \cdot \binom{n}{k}^2 \cdot a_k$$

Po dosadení za a_k *dostávame*

$D_{i_{n,k}} \leq 2^{n-k}\binom{n}{k}^2((p_1+p_3)^{2^k}-(p_1+p_3)^{2^{k+1}}-2p_3^{2^k}+2p_3^{2^k}(p_1+p_3)^{2^k}+p_3^{2^k}-p_3^{2^{k+1}}) = 2^{n-k} \cdot \binom{n}{k}^2 \cdot ((p_1+p_3)^{2^k}-p_3^{2^k}-((p_1+p_3)^{2^k}-p_3^{2^k})^2) \leq$

$$\leq 2^{n-k} \cdot \binom{n}{k}^2 \cdot ((p_1+p_3)^{2^k} - p_3^{2^k})$$

V nasledujúcom kroku využijeme Čebyševovu nerovnosť na ohraničenie strednej hodnoty náhodnej premennej $i_{n,k}$. Zvolíme
$b = \varphi(n)\binom{n}{k}\sqrt{2^{n-k} \cdot ((p_1+p_3)^{2^k} - p_3^{2^k})}$, kde $\frac{1}{\varphi(n)} = o(1)$.
Dostaneme

$$P[|i_{n,k} - E_{i_{n,k}}| \geq b] \leq \frac{D_{i_{n,k}}}{b^2} = \frac{1}{\varphi^2(n)} \to 0$$

Preto limita $P[|i_{n,k} - E_{i_{n,k}}| < b] = 1$. Na základe tohto ohraničenia môžeme sformulovať nasledujúcu vetu.

Veta 3.2.3 *S pravdepodobnosťou idúcej k 1 pre $n \to \infty$, pre všetky čiastočné boolovské funkcie f platí*

$$\binom{n}{k}\left(2^{n-k} \cdot ((p_1+p_3)^{2^k} - p_3^{2^k}) - \varphi(n)\sqrt{2^{n-k} \cdot ((p_1+p_3)^{2^k} - p_3^{2^k})}\right) < i_{n,k} <$$

$$< \binom{n}{k}\left(2^{n-k} \cdot ((p_1+p_3)^{2^k} - p_3^{2^k}) + \varphi(n)\sqrt{2^{n-k} \cdot ((p_1+p_3)^{2^k} - p_3^{2^k})}\right)$$

kde $\lim_{n\to\infty} \varphi(n) = \infty$.

Predpokladáme, že postupnosť (p_1+p_3) konverguje a $\frac{1}{p_1} = o(n) = \frac{1}{1-(p_1+p_3)} = \frac{1}{p_2}$. Predchádzajúca veta nám dáva možnosť získať asymptotické odhady $i_{n,k}$ pre k ohraničené v rámci istých limít. Využijeme to v nasledujúcej vete.

Veta 3.2.4 *Predpokladajme, že $\frac{1}{p_1} = o(n) = \frac{1}{p_2}$. S pravdepodobnosťou idúcou k 1 platí:*

(i) náhodná čiastočná boolovská funkcia neobsahuje intervaly rozmeru väčšieho ako $\mu = \lfloor \log n - \log\log \frac{1}{(p_1+p_3)} \rfloor + 1$. Naviac,

(ii) pre $k \leq \lfloor \log n - \log\log \frac{1}{(p_1+p_3)} \rfloor - 1 = \mu - 2$ počet $i_{n,k}(f)$ k-rozmerných intervalov čiastočnej boolovskej funkcie je asymptoticky rovný $\binom{n}{k}2^{n-k}((p_1+p_3)^{2^k} - p_3^{2^k})$, teda

$$i_{n,k} \sim \binom{n}{k} 2^{n-k}((p_1+p_3)^{2^k} - p_3^{2^k})$$

Dôkaz 3.2.3 *Časť (i). Nech $k > \mu$. Zvolíme $k = \mu + r$, kde $r \geq 1$. Ukážeme, že pre takto zvolené k horný odhad náhodnej premennej $i_{n,k}$ ide k 0 pre $n \to \infty$.*

Chceme ukázať, že:

$$\lim_{n\to\infty} \binom{n}{\mu+r} 2^{n-\mu-r}((p_1+p_3)^{2^{\mu+r}} - p_3^{2^{\mu+r}}) +$$

$$+\binom{n}{\mu+r}\varphi(n)\sqrt{2^{n-\mu-r}((p_1+p_3)^{2^{\mu+r}} - p_3^{2^{\mu+r)}}} = 0$$

$$\lim_{n\to\infty} \binom{n}{\mu+r} 2^{n-\mu-r}((p_1+p_3)^{2^{\mu+r}} - p_3^{2^{\mu+r}}) +$$

$$+ \lim_{n\to\infty}\binom{n}{\mu+r}\varphi(n)\sqrt{2^{n-\mu-r}((p_1+p_3)^{2^{\mu+r}} - p_3^{2^{\mu+r}})} = 0$$

Na to, aby sme dokázali že obidve limity idú k 0 nám stačí ukázať, že

$$\lim_{n\to\infty} \binom{n}{\mu+r}^2 \varphi^2(n) 2^{n-\mu-r}((p_1+p_3)^{2^{\mu+r}} - p_3^{2^{\mu+r}}) = 0$$

Platí:

$$\binom{n}{\mu+r}^2 \varphi^2(n) 2^{n-\mu-r}((p_1+p_3)^{2^{\mu+r}} - p_3^{2^{\mu+r}}) \leq$$

$$\leq \binom{n}{\mu+r}^2 \varphi^2(n) 2^{n-\mu-r}(p_1+p_3)^{2^{\mu+r}} \leq$$

$$\leq \varphi^2(n) 2^{2(\mu+r)\cdot\log n} 2^n 2^{2^{\mu+r}\cdot\log(p_1+p_3)} =$$

$$= \varphi^2(n) \frac{2^{2(\log n - \log\log\frac{1}{(p_1+p_3)}+1+r)\cdot\log n}}{2^{-n}2^{-2^{\log n-\log\log\frac{1}{(p_1+p_3)}+1+r}\cdot\log(p_1+p_3)}} =$$

V odhade použijeme nasledujúci výpočet $2^{\log n - \log\log\frac{1}{(p_1+p_3)}+1+r} = n \cdot \frac{1}{-\log(p_1+p_3)} \cdot 2 \cdot 2^r$

$$= \varphi^2(n) \frac{2^{2(\log n - \log\log\frac{1}{(p_1+p_3)}+1+r)\cdot\log n}}{2^{-n}2^{-n\cdot\frac{1}{-\log(p_1+p_3)}\cdot 2\cdot 2^r\cdot\log(p_1+p_3)}} =$$

$$= \varphi^2(n)\frac{2^{2(\log n-\log\log\frac{1}{(p_1+p_3)}+1+r)\cdot\log n}}{2^{(2^{r+1}-1)\cdot n}} \leq$$

$$\leq \varphi^2(n)\frac{2^{2(\log n-\log\log\frac{1}{p_1+p_3}+r+1)\log n}}{2^{(2^r-1)n}} = X_1(n)$$

Riešenie si môžeme rozdeliť na tri prípady:

1. $\lim_{n\to\infty} p_1+p_3 = p \in (0,1)$

$$\Rightarrow \lim_{n\to\infty} X_1(n) = 0$$

2. $\lim_{n\to\infty} p_1+p_3 = 0$, *potom* $\lim_{n\to\infty} \log\log\frac{1}{p_1+p_3} = \infty$

$$\Rightarrow \lim_{n\to\infty} X_1(n) = 0$$

3. $\lim_{n\to\infty} p_1+p_3 = 1$, *potom* $\lim_{n\to\infty} \log\log\frac{1}{p_1+p_3} = -\infty$, *avšak*

$$\lim_{n\to\infty} \frac{-\log\log\frac{1}{p_1+p_3}\log n}{n} = \lim_{n\to\infty}\frac{1}{n}\log\frac{1}{\log\frac{1}{p_1+p_3}}\log n =$$

$$= \lim_{n\to\infty}\frac{1}{n}\log(\frac{1}{\log(1+\frac{1-(p_1+p_3)}{p_1+p_3})})\log n = \lim_{n\to\infty}\frac{1}{n}\log\frac{1}{1-(p_1+p_3)}\log n = 0$$

$$\Rightarrow \lim_{n\to\infty} X_1(n) = 0$$

dokonca aj keď $\frac{1}{1-(p_1+p_3)} = o(n)$.

Časť (ii). Potrebujeme ukázať, že

$$\varphi(n)\binom{n}{k}\sqrt{2^{n-k}((p_1+p_3)^{2^k}-p_3^{2^k})} = o\left(\binom{n}{k}2^{n-k}((p_1+p_3)^{2^k}-p_3^{2^k})\right)$$

Teda, že

$$\lim_{n\to\infty}\frac{\varphi(n)}{\sqrt{2^{n-k}((p_1+p_3)^{2^k}-p_3^{2^k})}} = 0$$

pre vhodne zvolené $\varphi(n)$ a $k \leq \mu - 2$. Na to stačí overiť, či

$$\lim_{n\to\infty} 2^{n-k}((p_1+p_3)^{2^k} - p_3^{2^k}) = \infty$$

Použijeme binomickú vetu na to aby sme ohraničili tento výraz

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

$$\geq 2^{n-k}\frac{1}{n}(p_1+p_3)^{2^k} \geq 2^{n-\log n+\log\log\frac{1}{p_1+p_3}+1}\cdot 2^{-\log n}\cdot(p_1+p_3)^{2^{\log n-\log\log\frac{1}{p_1+p_3}-1}} =$$

$$= 2^{n-2\log n+\log\log\frac{1}{p_1+p_3}+1}\cdot 2^{n\cdot\frac{1}{-\log p_1+p_3}\cdot\frac{1}{2}\cdot\log p_1+p_3} =$$

$$= 2^{\frac{1}{2}n-2\log n+\log\log\frac{1}{p_1+p_3}+1} = X_2(n)$$

Na dôkaz toho, že $\lim_{n\to\infty} X_2(n) = \infty$ použijeme podobné výpočty ako v dôkaze prvej časti vety.

Keďže dĺžka úplnej DNF čiastočnej boolovskej funkcie je rovná $|N_f| = i_{n,0}(f)$, môžeme použiť predchádzajúcu vetu na asymptotický odhad počtu n-tíc, ktoré sa zobrazia na hodnotu 1.

Veta 3.2.5 *S pravdepodobnosťou idúcej k 1 pre $n \to \infty$, počet bodov patriacich do množiny N_f (dĺžka úplnej DNF čiastočnej boolovskej funkcie) je asymptoticky rovná $2^n \cdot p_1$*

$$|N_f| = i_{n,0}(f) \sim 2^n \cdot ((p_1+p_3)^{2^0} - p_3^{2^0}) = 2^n \cdot p_1$$

3.3 Maximálne intervaly

V tejto kapitole sa budeme snažiť odhadnúť zložitosť skrátenej DNF čiastočnej boolovskej funkcie, ktorá je rovná počtu všetkých maximálnych intervalov. Na to potrebujeme určiť strednú hodnotu počtu k-rozmerných maximálnych intervalov funkcie.

Nech $I_{n,k}$ je počet k-rozmerných maximálnych intervalov funkcie $f \in P^n$ a $s(f)$ je zložitosť skrátenej DNF funkcie f. Potom $s(f) = \sum_{k=0}^{n} I_{n,k}(f)$.

Veta 3.3.1

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

Dôkaz 3.3.1 *Nech K je ľubovoľná k-rozmerná podkocka kocky B^n. Definujeme indikátor θ_n nasledovne*

$$\theta_K(f) = \begin{cases} 1 & \text{ak } K \text{ je maximálny interval funkcie } f \\ 0 & \text{inak} \end{cases}$$

$$f \in B_n$$

Teraz pre náhodnú premennú platí $I_{n,k} = \sum_K \theta_K$, kde K ide cez všetky k-rozmerné podkocky B_n. Preto stačí vypočítať $EI_{n,k} = P[\theta_K(f) = 1]$ pre pevne zvolené K.

Dekompozíciou čiastočnej boolovskej funkcie f vzhľadom na prvých $n-k$ premenných dostávame

$$f(x_1,\ldots,x_n) = \bigvee_{\sigma=(\sigma_1,\ldots,\sigma_{n-k})\in B_{n-k}} x_1^{\sigma_1}\wedge x_2^{\sigma_2}\wedge\ldots\wedge x_{n-k}^{\sigma_{n-k}}\wedge f(\sigma, x_{n-k+1},\ldots,x_n) =$$

$$= x_1\wedge x_2\wedge\ldots\wedge x_{n-k}\wedge f_0(x_{n-k+1},\ldots,x_n)\vee\bar{x}_1\wedge x_2\wedge\ldots\wedge x_{n-k}\wedge f_1(x_{n-k+1},\ldots,x_n)\vee\ldots$$

$$\ldots\vee x_1\wedge x_2\wedge\ldots\wedge\bar{x}_{n-k}\wedge f_{n-k}(x_{n-k+1},\ldots,x_n)\vee\ldots$$

Podkocka K je maximálnym intervalom práve vtedy, keď neexistuje podkocka dimenzie $k+1$, ktorá by ju obsahovala. Inými slovami, keď podkocku K nemožeme rozšíriť žiadnym z $n-k$ smerov, spojiť ju s podkockou dimenzie k a vytvoriť podkocku dimenzie $k+1$. Teda K je maximálnym intervalom práve vtedy, keď platí

(i) f_0 *sa rovná 1 alebo nie je def. a obsahuje aspoň jednu 1*

(ii) f_i *obsahuje aspoň jednu 0 pre* $i = 1, 2, \ldots, n-k$

Dostávame

$$P[f \in B_n; (i)] = (p_1 + p_3)^{2^k} - p_3^{2^k}$$

$$P[f \in B_n; (ii)] = 1 - (p_1 + p_3)^{2^k}$$

pre pevne zvolené n.

Pretože udalosti f_i *a* f_j *sú nezávislé pre* $i \neq j$, *pravdepodobnosť* $P[\theta_K(f) = 1] = ((p_1 + p_3)^{2^k} - p_3^{2^k})(1 - (p_1 + p_3)^{2^k})^{n-k}$. *Keďže v* B_n *existuje* $\binom{n}{k}2_{n-k}$ *k-rozmerných podkociek, dostávame*

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

V nasledujúcich vetách odhadneme zložitosť skrátenej DNF náhodnej čiastočnej boolovskej funkcie.

Veta 3.3.2 $Es \leq 2^n \cdot n^{(1+\varepsilon_1(n))\cdot \log\log \frac{1}{p_1+p_3}\, n}$, $\varepsilon_1(n) \to 0$. *Navyše, ak* $\lim_{n\to\infty} p_1 + p_3 = p \in (0,1)$, *potom* $\varepsilon_1(n) = \mathrm{O}(\frac{1}{\log\log \frac{1}{p_1+p_3}\, \mathrm{n}})$.

Dôkaz 3.3.2 *Vieme už, že* $s(f) = \sum_{k=0}^{n} I_{n,k}$. *Najprv nás bude zaujímať* $\max_{0\leq k\leq n} EI_{n,k}$. *Pozrieme sa na pomer*

$$\frac{EI_{n,k+1}}{EI_{n,k}} = \frac{\binom{n}{k+1}2^{n-k-1}((p_1 + p_3)^{2^{k+1}} - p_3^{2^{k+1}})(1 - (p_1 + p_3)^{2^{k+1}})^{n-k-1}}{\binom{n}{k}2^{n-k}((p_1 + p_3)^{2^k} - p_3^{2^k})(1 - (p_1 + p_3)^{2^k})^{n-k}} =$$

$$= \frac{(n-k)}{2(k+1)} \cdot \frac{(1 + (p_1 + p_3)^{2^k})^{n-k-1}}{\frac{1+p_3^{2^k}}{(p_1+p_3)^{2^k}+p_3^{2^k}} - 1} = X_3(n,k)$$

Pri ohraničovaní $X_3(n,k)$ *použijeme nasledujúce nerovnosti.*

Pre $k \leq \log\log_{\frac{1}{p_1+p_3}} n - 1 = \log\log n - \log\log\frac{1}{p_1+p_3} - 1$ *platí*

$$(p_1+p_3)^{2^k} \geq (p_1+p_3)^{2^{\log\log_{\frac{1}{p_1+p_3}} n - 1}} = \left((p_1+p_3)^{2^{\log\log_{\frac{1}{p_1+p_3}} n}}\right)^{\frac{1}{2}} = \frac{1}{\sqrt{n}}$$

Pre $k > \log\log n - \log\log\frac{1}{p_1+p_3}$ *platí*

$$(p_1+p_3)^{2^k} < (p_1+p_3)^{2^{\log\log_{\frac{1}{p_1+p_3}} n}} = \frac{1}{n}$$

Potrebujeme ukázať, že pre $k \leq \log\log n - \log\log\frac{1}{p_1+p_3} - 1$ *a dostatočne veľké* n *platí* $X_3(n,k) > 1$.

$$X_3(n,k) \geq \frac{n - \log\log n + \log\log\frac{1}{p_1+p_3} + 1}{2(\log\log n - \log\log\frac{1}{(p_1+p_3)})} \cdot \frac{1}{\sqrt{n}} \cdot \left(1+\frac{1}{\sqrt{n}}\right)^{n-\log\log n+\log\log\frac{1}{p_1+p_3}}$$

Môžu nastať tri prípady:

1. $\lim_{n\to\infty} p_1 + p_3 = p \in (0,1)$
2. $\lim_{n\to\infty} p_1 + p_3 = 0$
3. $\lim_{n\to\infty} p_1 + p_3 = 1$

Na overenie všetkých troch prípadov použijeme rovnaké argumenty ako v predchádzajúcich dôkazoch.

$$\lim_{n\to\infty} \frac{n - \log\log n + \log\log\frac{1}{p_1+p_3} + 1}{2(\log\log n - \log\log\frac{1}{(p_1+p_3)})\frac{1}{\sqrt{n}}} = \infty = \lim_{n\to\infty} \left(1+\frac{1}{\sqrt{n}}\right)^{n-\log\log n+\log\log\frac{1}{p_1+p_3}}$$

Preto $\lim_{n\to\infty} X_3(n,k) = \infty$ *pre* $k \leq \log\log n - \log\log\frac{1}{p_1+p_3} - 1$. *Z toho vyplýva, že pomer* $\frac{EI_{n,k+1}}{EI_{n,k}} > 1$ *pre dostatočne veľké* n.

Ak vezmeme $k \geq \log\log n - \log\log \frac{1}{p_1+p_3}$,

$$0 \leq X_3(n,k) \leq \frac{n - \log\log n + \log\log \frac{1}{p_1+p_3}}{2(\log\log n - \log\log \frac{1}{(p_1+p_3)} + 1)} \cdot \frac{1}{n-1} \cdot \left(1+\frac{1}{n}\right)^{n-\log\log n+\log\log \frac{1}{p_1+p_3}-1}$$

Dostávame limity

$$\lim_{n\to\infty} \frac{n - \log\log n + \log\log \frac{1}{p_1+p_3}}{2(\log\log n - \log\log \frac{1}{(p_1+p_3)} + 1) \cdot (n-1)} = 0$$

$$\lim_{n\to\infty} \left(1+\frac{1}{n}\right)^{n-\log\log n+\log\log \frac{1}{p_1+p_3}-1} = e$$

Preto $\lim_{n\to\infty} X_3(n,k) = 0$ *pre* $k > \log\log n - \log\log \frac{1}{p_1+p_3}$ *a* $\frac{EI_{n,k+1}}{EI_{n,k}} < 1$ *pre dostatočne veľké n.*

Z toho vyplýva, že $EI_{n,k}$ *ako funkcia od k nadobúda svoje maximum buď pre k rovné* $\lfloor \log\log n - \log\log \frac{1}{p_1+p_3} \rfloor = \lambda$ *alebo* $\lfloor \log\log n - \log\log \frac{1}{p_1+p_3} \rfloor + 1 = \lambda + 1$. *Naviac,*

$$\max_k EI_{n,k} \leq \binom{n}{\lambda+1} \cdot 2^{n-\lambda} \cdot ((p_1+p_3)^{2^\lambda} - p_3^{2^{\lambda+1}}) \cdot (1-(p_1+p_3)^{2^{\lambda+1}})^{n-\lambda} \leq$$

$$\leq n^{\lambda+1} \cdot 2^{n-\lambda} \cdot (p_1+p_3)^{2^\lambda} \cdot (1-(p_1+p_3)^{2^{\lambda+1}})^{n-\lambda} \leq$$

$$\leq n^{\lambda+1} \cdot 2^{n-\lambda} \cdot \frac{1}{n} \cdot (1-\frac{1}{n^2})^{n-\lambda} \leq$$

$$\leq 2^n \cdot n^{(1+\varepsilon_1'(n)) \cdot \log\log \frac{1}{p_1+p_3} n}$$

kde $\varepsilon_1'(n) = \frac{1+\frac{1}{2}\log n - \log\log \frac{1}{p_1+p_3} n}{\log n \cdot \log\log \frac{1}{p_1+p_3} n} \to 0$. *Ak* $\lim_{n\to\infty} p_1 + p_3 = p \in (0,1)$, *potom* $\varepsilon_1'(n) = \mathrm{O}(\frac{1}{\log\log \frac{1}{p_1+p_3} \mathrm{n}})$.

Zhrnutím týchto výsledkov získavame nasledujúci odhad

$$Es \leq (n+1) \cdot \max_k EI_{n,k} \leq$$

$$\leq n^{\frac{\log(n+1)}{\log n}} \cdot 2^n \cdot n^{(1+\varepsilon_1'(n))\cdot\log\log\frac{1}{p_1+p_3}\,n} =$$
$$= 2^n \cdot n^{(1+\varepsilon_1(n))\cdot\log\log\frac{1}{p_1+p_3}\,n}$$

Veta 3.3.3 *S pravdepodobnosťou idúcou k 1 pre* $n \to \infty$ *platí, že* $s(f) \geq \frac{i_{n,\lambda}(f)}{\binom{\mu}{\lambda}2^{\mu-\lambda}}$, *kde* $\mu = \lfloor \log n - \log\log\frac{1}{p_1+p_3} \rfloor + 1$ *a* $\lambda = \lfloor \log\log n - \log\log\frac{1}{p_1+p_3} \rfloor + 1$.

Dôkaz 3.3.3 *Vieme už, že čiastočná boolovská funkcia f neobsahuje maximálne intervaly dimenzie väčšej ako* μ *takmer určite. Z toho vyplýva, že pre každý maximálny interval platí, že neobsahuje viac ako* $\binom{\mu}{\lambda}2^{\mu-\lambda}$ λ*-rozmerných intervalov. Na druhej strane, každý* λ*-rozmerný podinterval je obsiahnutý v aspoň jednom maximálnom intervale funkcie f. Preto*

$$i_{n,\lambda}(f) \leq s(f) \cdot \binom{\mu}{\lambda} 2^{\mu-\lambda}$$

$$s(f) \geq \frac{i_{n,\lambda}(f)}{\binom{\mu}{\lambda}2^{\mu-\lambda}}$$

takmer určite.

V nasledujúcej časti skombinujeme predchádzajúce vety, aby sme získali odhad náhodnej premennej $s(f)$, teda zložitosti skrátenej DNF čiastočnej boolovskej funkcie f. Dokázali sme, že $s(f) \leq U.B.$, $s(f) \geq L.B.$, teda hornú a dolnú hranicu $s(f)$ takmer určite. My chceme ale dokázať ohraničenie $L.B. \leq s(f) \leq U.B.$ takmer určite. Označíme si množiny

$$F_n = \{f \in P^n | s(f) \leq U.B.\}$$

$$G_n = \{f \in P^n | s(f) \geq L.B.\}$$

Potom

$$\lim_{n\to\infty} P[F_n] = 1 = \lim_{n\to\infty} P[G_n]$$

Teraz

$$1 \geq \lim_{n\to\infty} P[F_n \cap G_n] = 1 - \lim_{n\to\infty} P[P^n - (F_n \cap G_n)] \geq$$
$$\geq 1 - \lim_{n\to\infty} P[P^n - F_n] - \lim_{n\to\infty} P[P^n - G_n] = 1$$

Ale $\lim_{n\to\infty} P[F_n \cap G_n] = 1$ je ekvivalentná tvrdeniu $L.B. \leq s(f) \leq U.B.$ takmer určite.

Takéto kombinovanie takmer istých udalostí nám umožňuje lema, ktorá dovoľuje obmedziť naše konštrukcie na podmnožiny P^n, ktorých pravdepodobnosť pre $n \to \infty$ ide k 1.

Lema 3.3.0.4 *Nech $(X_n, S_n, Q_n)_{n=1}^{\infty}$ je postupnosť pravdepodobnostných priestorov. Predpokladajme, že $Y_n, Z_n \in S_n$ sú udalosti také, že $\lim_{n\to\infty} Q_n(Y_n) = 1 = \lim_{n\to\infty} Q_n(Z_n)$ a nech $A_n \in S$. Potom*

(i) $\lim_{n\to\infty} Q_n(Y_n \cap Z_n) = 1$

(ii) $\lim_{n\to\infty} Q_n(A_n|Y_n) = 1 \Rightarrow \lim_{n\to\infty} Q_n(A_n) = 1.$

Dôkaz 3.3.4 *Dôkaz je triviálny, z toho dôvodu ho nebudeme uvádzať.*

Veta 3.3.4 *Predpokladajme, že $\frac{1}{p_1} = o(n) = \frac{1}{p_2}$. Potom s pravdepodobnosťou idúcou k 1 pre $n \to \infty$ platí, že*

$$2^n \cdot n^{(1-\varepsilon_2(n))\cdot \log\log \frac{1}{p_1+p_3}\, n} \leq s(f) \leq 2^n \cdot n^{(1+\varepsilon_3(n))\cdot \log\log \frac{1}{p_1+p_3}\, n} \qquad \varepsilon_2(n), \varepsilon_3(n) \to 0$$

Ak $\lim_{n\to\infty} p_1 + p_3 = p \in (0,1)$, *potom* $\varepsilon_2(n) = \mathrm{O}(\frac{1}{\log\log \frac{1}{p_1+p_3}\, \mathrm{n}}) = \varepsilon_3(\mathrm{n})$

Dôkaz 3.3.5 *Na dôkaz horného ohraničenia použijeme Markovovu nerovnosť s $b = n$, ktorú potom spojíme s Vetou 3.3.2 o odhade strednej hodnoty Es. Dostávame*

$$s(f) \leq n \cdot Es \qquad \textit{takmer určite}$$

$$s(f) \leq 2^n \cdot n^{(1+\varepsilon_3(n))\cdot \log\log_{\frac{1}{p_1+p_3}} n} \qquad kde\ \ \varepsilon_3(n) = \varepsilon_1(n) + \frac{1}{\log\log_{\frac{1}{p_1+p_3}} n} \to 0$$

Platí $\varepsilon_3(n) = \mathrm{O}(\frac{1}{\log\log_{\frac{1}{p_1+p_3}} \mathrm{n}})$ *keď* $\lim_{n\to\infty} p_1 + p_3 = p \in (0,1)$.

Pri dôkaze dolného ohraničenia najskôr odhadneme náhodnú premennú $i_{n,\lambda}(f)$. *Z Vety 3.2.3 vyplýva, že*

$$i_{n,\lambda}(f) \geq \binom{n}{\lambda}\left(2^{n-\lambda}\cdot((p_1+p_3)^{2^\lambda} - p_3^{2^\lambda}) - \varphi(n)\sqrt{2^{n-\lambda}\cdot((p_1+p_3)^{2^\lambda} - p_3^{2^\lambda})}\right)$$

Použitím nerovnosti $\binom{n}{\lambda} > (\frac{n}{\lambda})^\lambda$, *vlastnosti koeficientov v binomickej vete a substitúciov* λ *dostaneme*

$$\begin{aligned} i_{n,\lambda}(f) &\geq \left(\frac{n}{\lambda}\right)^\lambda\left(2^{n-\lambda}\cdot\frac{p_1}{p_1+p_3}\cdot(p_1+p_3)^{2^\lambda} - \varphi(n)\cdot\sqrt{2^{n-\lambda}\cdot(p_1+p_3)^{2^\lambda}}\right) \geq \\ &\geq 2^n\cdot n^\lambda\cdot\lambda^{-\lambda}\cdot(2^{-\lambda}\cdot n^{-2} - n\cdot 2^{\frac{-n-\lambda}{2}}\cdot n^{-\frac{1}{2}}) \geq \\ &\geq 2^n\cdot n^\lambda\cdot n^{\frac{-\lambda\log\lambda}{\log n}}\cdot n^{-2}\cdot 2^{-\lambda}\cdot(1 - n^{\frac{5}{2}}\cdot 2^{\frac{\lambda-n}{2}}) > \\ &> 2^n\cdot n^{\lambda - \frac{\lambda\log\lambda}{\log n} - 3}\cdot(1 - n^{\frac{5}{2}}\cdot 2^{\frac{\lambda-n}{2}}) > \\ &> 2^n\cdot n^{(1-\varepsilon_2'(n))\cdot\log\log_{\frac{1}{p_1+p_3}} n} \end{aligned}$$

$$kde\ \varepsilon_2'(n) = \frac{3\log n + \log\log_{\frac{1}{p_1+p_3}} n\cdot\log\log\log_{\frac{1}{p_1+p_3}} n - \log\left(1-\sqrt{\frac{n^5}{2^n}\cdot\log_{\frac{1}{p_1+p_3}} n}\right)}{\log n\cdot\log\log_{\frac{1}{p_1+p_3}} n} \to 0$$

Aplikovaním Vety 3.3.3 získame

$$s(f) \geq \frac{i_{n,\lambda}}{\binom{\mu}{\lambda}2^{\mu-\lambda}} \geq \frac{2^n\cdot n^{(1-\varepsilon_2'(n))\cdot\log\log_{\frac{1}{p_1+p_3}} n}}{\mu^\lambda\cdot 2^{\mu-\lambda}} \geq 2^n\cdot n^{(1-\varepsilon_2(n))\cdot\log\log_{\frac{1}{p_1+p_3}} n}$$

$$kde\ \varepsilon_2(n) = \varepsilon_2'(n) + \frac{\log(\log n - \log\log\frac{1}{p_1+p_3}+1)}{\log n} \to 0$$

Ďalšie výpočty ukazujú, že $\varepsilon_2(n)$ *je rádu* $\mathrm{O}(\frac{1}{\log\log \frac{1}{p_1+p_3} n})$ *práve vtedy, keď* $\lim_{n\to\infty} p_1 + p_3 = p \in (0,1)$ *a je rôzna od 0 a 1.*

Záver

V závere práce niekoľko poznámok. Predtým, ako sme sa pokúsili odhadnúť zložitosti DNF čiastočnej boolovskej funkcie sme si definovali pravdepodobnostný model. Na začiatku sme sa pokúsili pracovať s modelom, v ktorom bol z 2^n vrcholov presne daný počet bodov, v ktorých náhodná čiastočná boolovská funkcia nie je definovaná a pravdepodobnosť $p = \frac{1}{2}$ s ktorou sa hodnota bodu, v ktorom je funkcia definovaná zobrazí na hodnotu 0 resp. 1. V tomto prípade sme sa dostali k zložitým sumám, ktoré bolo technicky náročné zvládnuť a vyjadriť ich v kompaktnom tvare. Preto sme sa rozhodli pre iný pravdepodobnostný model. Máme teda pravdepodobnosť p_1, s ktorou sa vrchol zobrazí na hodnotu 1, pravdepodobnosť p_2, s ktorou sa vrchol zobrazí na hodnotu 0 a pravdepodobnosť p_3, s ktorou funkcia nie je v danom vrchole definovaná. Pre takýto pravdepodobnostný model sme získali horné a dolné ohraničenie, strednú hodnotu a disperziu pre náhodnú premennú - počet k-rozmerných podkociek čiastočnej boolovskej funkcie a asymptotický odhad tejto premennej, tiež sme ohraničili k, teda dimenziu týchto podkociek. Na základe týchto výsledkov sme odhadli zložitosť úplnej DNF funkcie. V druhej časti sme získali strednú hodnotu pre počet maximálnych intervalov a dolný a horný odhad pre zložitosť dĺžky skrátenej DNF čiastočnej boolovskej funkcie. V prípade čiastočných náhodných boolovských funkcií je možné formulovať zmysluplné problémy.

Literatúra

[1] Sergej Vsevolovic Yablonski, Úvod do diskrétnej matematiky, Alfa, Vydavateľstvo technickej a ekonomickej literatúry, Bratislava, 1984.

[2] M. Škoviera, On the Minimization of Random boolean Functions, Part 1, Computers and Artificial Intelligence, Bratislava, 1986.

[3] M. Škoviera, On the Minimization of Random boolean Functions, Part 2, Computers and Artificial Intelligence, Bratislava, 1986.

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

[5] Lin Sin-Ljan, O sravnenii složnostej minimaľnych i kratčajších dizjunktivnych normáľnych form dlja funkciij algebry logiki, Sb. Problemy kibernetiki, M., Nauka, 1967, vyp. 18, s. 11

[6] Glagoljev V. V, O dline tupikuvoj dizjunktivnoj normaľnoj formy. mat., Zametky, 2. N^o6, 665, (1967)

[7] Vasiliev L. V., O sravnenii složnosti tupikovych i minimaľnych dizjunktivnych normalnych form, Sb. Problemy kibernetiki, M., Fizmatgiz, 1963, vyp. 10, s. 5

[8] Sergej Vsevolovic Yablonski and Lupanov O. B., Discrete Mathematics and Mathematical Problems of Cybernetics, Nauka, Moscow, 1974 (in Russian)

[9] C. M. Grinstead and J. L. Snell, Introduction to Probability, Amer Mathematical Society, 1997