1 Fourierova transformace Booleovských funkcí

Kromě proměnných $x_i \in \{0,1\}$, zavedeme ještě proměnné $z_i = (-1)^{x_i}$. Pro množinu indexů $I \subseteq \{1,\ldots,n\}$ nechť

$$Z_I = \prod_{i \in I} z_i = (-1)^{\oplus_{i \in I} x_i} \ ,$$

speciálně $Z_\emptyset = 1$. Pro Booleovskou funkci $g : \{0,1\}^n \to \{0,1\}$ budeme značením $g' : \{1,-1\}^n \to \{1,-1\}$ rozumět funkci splňující

$$g'(z_1,\ldots,z_n) = (-1)^{g(x_1,\ldots,x_n)}$$

při výše uvedené transformaci proměnných a tuto funkci budeme považovat za prvek lineárního prostoru funkcí $\{1,-1\}^n \to \mathrm{R}$, který budeme značit V_n. Pro funkce $g_1, g_2 : \{1,-1\}^n \to \mathrm{R}$ definujme skalární součin jako

$$\langle g_1, g_2 \rangle = \frac{1}{2^n} \sum_{z \in \{1,-1\}^n} g_1(z) g_2(z) \ .$$

Věta 1.1 *Funkce Z_I tvoří ortonormální bázi prostoru funkcí V_n.*

Jako důsledek dostaneme, že každou funkci z V_n lze jednoznačně vyjádřit jako lineární kombinaci Z_I, tedy jako multilineární polynom proměnných z_i.

Jestliže je funkce $g(x_1,\ldots,x_n)$ vyjádřena jako polynom s proměnnými x_i nad reálnými čísly, tedy jako lineární kombinace

$$X_I = \prod_{i \in I} x_i \ ,$$

můžeme získat g' v podobě polynomu s proměnnými z_i tak, že ve vyjádření $g' = 1 - 2g(x_1,\ldots,x_n)$ dosadíme za x_i podle následujících transformací

$$\begin{aligned} z_i &= 1 - 2x_i \ , \\ x_i &= \frac{1 - z_i}{2} \ . \end{aligned}$$

Protože tyto transformace jsou lineární, mají funkce g a g' vyjádření pomocí polynomu stejného stupně.

Příklad. Uvažme funkci

$$and(x_1, x_2) = x_1 x_2 \ .$$

Pak

$$and'(z_1, z_2) = 1 - 2 \cdot \frac{1 - z_1}{2} \cdot \frac{1 - z_2}{2} \ ,$$

tedy po úpravě

$$and'(z_1, z_2) = \frac{1}{2}(1 + z_1 + z_2 - z_1 z_2) \ .$$

Protože negace znamená pro proměnné z_i změnu znaménka, dostaneme

$$or'(z_1, z_2) = -and'(-z_1, -z_2) = -1 + 2 \cdot \frac{1 + z_1}{2} \cdot \frac{1 + z_2}{2} \ ,$$

tedy po úpravě

$$or'(z_1, z_2) = \frac{1}{2}(-1 + z_1 + z_2 + z_1 z_2) \ .$$

Pro koeficienty vyjádření funkce $g \in V_n$ pomocí Z_I platí následující.

Věta 1.2 *Jestliže $g \in V_n$, pak pro každé $z \in \{1, -1\}^n$ platí*

$$g(z) = \sum_{I \subseteq \{1,\ldots,n\}} a_I Z_I$$

právě tehdy, když pro každou množinu indexů $I \subseteq \{1, \ldots, n\}$ je $a_I = \langle g, Z_I \rangle$.

Speciálně, pokud g reprezentuje nějakou Booleovskou funkci, pak platí

Věta 1.3 *Jestliže $g : \{1, -1\}^n \to \{1, -1\}$, pak pro koeficienty a_I polynomu, který reprezentuje g, platí*

$$\sum_{I \subseteq \{1,\ldots,n\}} a_I^2 = 1 \ .$$

Důkaz. Obecněji, v prostoru V_n platí pro koeficienty a_I Parsevalova rovnost

$$\sum_{I \subseteq \{1,\ldots,n\}} a_I^2 = \langle g, g \rangle \ .$$

Protože g nabývá pouze hodnot $1, -1$, platí $\langle g, g \rangle \ = 1$. □

Z Věty 1.3 plyne, že koeficienty a_I v polynomech, které vyjadřují Booleovské funkce, jsou v absolutní hodnotě nejvýše 1. Navíc, pokud má některý koeficient absolutní hodnotu 1, jsou ostatní koeficienty rovny 0. Tato situace nastává právě pro lineární Booleovské funkce, protože ty jsou reprezentovány polynomy Z_I a $-Z_I$ pro některou množinu $I \subseteq \{1, \ldots, n\}$.

Všechny koeficienty pro funkce and' a or' mají stejnou absolutní hodnotu. Z Věty 1.3 plyne, že pak musí mít absolutní hodnotu $1/2$. Existují i další funkce, které mají stejnou absolutní hodnotu všech koeficientů. Tyto funkce existují jen pro sudé n a společná absolutní hodnota koeficientů je $2^{-n/2}$. Příklady takových

funkcí lze získat násobením polynomů and' nebo or' pro disjunktní množiny proměnných. Dostaneme tak například funkci $kvadr'$, jejíž Booleovská forma je

$$kvadr(x_1, \ldots, x_{2k}) = \bigoplus_{i=1}^{k} x_{2i-1}x_{2i} \ .$$

Podobným výpočtem jako pro disjunkci dvou proměnných, dostaneme pro libovolné $n \geq 2$

$$or'(z_1, \ldots, z_n) = -1 + 2 \cdot \frac{1+z_1}{2} \cdot \ldots \cdot \frac{1+z_n}{2} \ ,$$

což po roznásobení dává

$$or'(z_1, \ldots, z_n) = -1 + \frac{1}{2^{n-1}} \sum_I Z_I \ ,$$

a po úpravě

$$or'(z_1, \ldots, z_n) = -\frac{2^{n-1}-1}{2^{n-1}} + \frac{1}{2^{n-1}} \sum_{I \neq \emptyset} Z_I \ .$$

Jestliže $sel(x_1, x_2, x_3)$ je definována vztahy $sel(0, x_2, x_3) = x_2$ a $sel(1, x_2, x_3) = x_3$, pak lze rozborem případů pro hodnotu z_1 ověřit, že

$$sel'(z_1, z_2, z_3) = \frac{1}{2}(z_2 + z_3 + z_1(z_2 - z_3)) \ .$$

Odvoďme vyjádření maj_3', kde funkcí maj_3 rozumíme

$$maj_3(x_1, x_2, x_3) = x_1x_2 \vee x_1x_3 \vee x_2x_3 \ .$$

Platí

$$maj_3(x_1, x_2, x_3) = sel(x_1, x_2x_3, x_2 \vee x_3) \ .$$

tedy

$$maj_3'(z_1, z_2, z_3) = sel'(z_1, and'(z_2, z_3), or'(z_2, z_3)) \ .$$

Dosazením polynomů pro and' a or' do polynomu pro sel' dostaneme

$$maj_3'(z_1, z_2, z_3) = \frac{1}{2}(z_1 + z_2 + z_3 - z_1z_2z_3) \ .$$

2 Dolní odhad složitosti pro rozhodovací stromy

Věta 2.1 *Nechť f je Booleovská funkce n proměnných a a_I jsou koeficienty polynomu pro f'. Pak pro každou množinu indexů $I \subseteq \{1, \ldots, n\}$ platí*

$$\mathrm{dt}(f) \geq 2^{|I|} \sum_{J \supseteq I} |a_J| \ .$$

Důkaz. Zvolme rozhodovací strom pro f velikosti $s = \mathrm{dt}(f)$ a předpokládejme, že listy jsou očíslovány $1, \ldots, s$. Pro $i = 1, \ldots, s$, nechť A_i je množina proměnných testovaných na cestě z kořene do i-tého listu, $v_i \in \{1, -1\}$ je hodnota funkce f', která je tomuto listu přiřazena, a g_i je charakteristická funkce množiny vstupů, pro které dojde výpočet do i-tého listu. Pak platí

$$f' = \sum_{i=1}^{s} v_i g_i \ .$$

Pro libovolnou $J \subseteq \{1, \ldots, n\}$ tedy platí

$$a_J = \langle f', Z_J \rangle = \sum_{i=1}^{s} v_i \langle g_i, Z_J \rangle \ .$$

Pokud $J \not\subseteq A_i$, pak $\langle g_i, Z_J \rangle = 0$, protože funkce Z_J není konstantní na podkrychli definované zafixováním proměnných z množiny A_i, a tedy nabývá obou možných hodnot na stejném počtu prvků této podkrychle. Tento fakt umožní úpravu sumačních indexů v následujícím odvození. Pro libovolnou $I \subseteq \{1, \ldots, n\}$ platí

$$\sum_{I \subseteq J} |a_J| \leq \sum_{I \subseteq J} \sum_{i=1}^{s} |\langle g_i, Z_J \rangle| = \sum_{i=1}^{s} \sum_{I \subseteq J \subseteq A_i} |\langle g_i, Z_J \rangle| \ .$$

Pro další úpravu využijeme toho, že pro $J \subseteq A_i$ je Z_J konstantní na podkrychli, kde jsou zafixovány proměnné z A_i a kde je $g_i = 1$. Pro $J \subseteq A_i$ tedy platí

$$|\langle g_i, Z_J \rangle| = \frac{1}{2^n} 2^{n - |A_i|} = 2^{-|A_i|} \ ,$$

což spolu s předchozím odvozením dává

$$\sum_{I \subseteq J} |a_J| \leq \sum_{i=1}^{s} \sum_{I \subseteq J \subseteq A_i} 2^{-|A_i|} = \sum_{i=1}^{s} 2^{|A_i| - |I|} \cdot 2^{-|A_i|} = s\, 2^{-|I|} \ .$$

Platí tedy

$$\mathrm{dt}(f) = s \geq 2^{|I|} \sum_{I \subseteq J} |a_J| \ ,$$

což jsme měli dokázat. □

Uvažme následující dvě posloupnosti Booleovských funkcí. Pro $h \geq 0$ nechť F_h je funkce počítaná formulí se spojkou maj_3, která má tvar vyváženého stromu hloubky h, tedy s 3^h listy, přičemž v každém listě je jiná vstupní proměnná. Podobně, nechť G_h je funkce počítaná formulí se spojkou NAND, která má tvar vyváženého stromu hloubky h, tedy s 2^h listy, které opět obsahují různé proměnné.

Věta 2.2 *Platí* $\mathrm{dt}(F_h) \geq 2^{\Omega(3^h)}$ *a* $\mathrm{dt}(G_h) \geq 2^{\Omega(2^h)}$.

Důkaz. Dokážeme jen tvrzení pro F_h. Nechť F'_h je funkce F_h v reprezentaci pomocí $\{1, -1\}$ a nechť b_h je koeficient u monomu nejvyššího stupně, tedy u součinu všech proměnných, v polynomu pro F'_h. Víme, že platí

$$F'_1 = maj'_3(z_1, z_2, z_3) = \frac{1}{2}(z_1 + z_2 + z_3 - z_1 z_2 z_3) \ .$$

a tedy $b_1 = -1/2$. Pro každé $h \geq 2$ platí $F'_h = maj'_3(F'_{h-1,1}, F'_{h-1,2}, F'_{h-1,3})$, kde $F'_{h-1,j}$ jsou funkce F'_{h-1} na disjunktních množinách proměnných. Je zřejmé, že do koeficientu nejvyššího stupně v F'_h přispívají pouze koeficienty u nejvyššího stupně v F'_{h-1}. Platí tedy rekurence

$$b_h = -\frac{1}{2} b_{h-1}^3 \ ,$$

ze které plyne

$$b_h = (-1)^h \cdot \left(\frac{1}{2}\right)^{(3^h-1)/2} \ .$$

Označme jako I množinu indexů všech proměnných F_h. Protože $|I| = 3^h$, dostaneme pomocí Věty 2.1 nerovnost

$$\mathrm{dt}(F_h) \geq 2^{|I|} |b_h| = 2^{(3^h+1)/2} \ .$$

Z toho již požadované tvrzení pro F_h plyne.

Pro funkci G_h lze použít podobný postup, ale odhaduje se součet absolutních hodnot všech koeficientů polynomu pro G'_h a Věta 2.1 se použije pro $I = \emptyset$. Důkaz dolního odhadu pro $\mathrm{dt}(G_h)$ nebudeme uvádět. □

3 Souvislost s velikostí DNF

Dolní odhady složitosti rozhodovacích stromů pro funkce F_h a G_h z předchozí sekce nyní dáme do souvislosti se složitostí DNF pro tyto funkce a jejich negace. Protože je funkce F_h monotonní, je její složitost DNF rovna počtu primárních implikantů. Funkce F_h je navíc samoduální, tedy složitost DNF pro F_h a $\neg F_h$ je stejná. Primární implikanty F_h odpovídají podstromům formule, které mají kořen shodný s celou formulí a v každé spojce pokračují právě do dvou jejích argumentů. Implikant je reprezentován dosazením hodnoty 1 do všech proměnných, které jsou v listech příslušného podstromu. Uvažované podstromy mají 2^h listů a $2^h - 1$ vnitřních uzlů. Pro výběr konkrétního podstromu je v každém jeho vnitřním uzlu třeba vybrat jednu ze 3 možných variant jeho následníků. Všech uvažovaných podstromů je tedy právě 3^{2^h-1} a platí tedy

$$\mathrm{dnf}(F_h) = \mathrm{dnf}(\neg F_h) = 3^{2^h-1} \ .$$

Jako důsledek odvodíme, že velikost rozhodovacího stromu pro F_h je quasi-polynomiálně větší než součet složitosti DNF pro F_h a její negaci. Platí následující tvrzení

Věta 3.1 *Pro libovolné h nechť $N_h = \mathrm{dnf}(F_h) + \mathrm{dnf}(\neg F_h)$. Pak platí*

$$\mathrm{dt}(F_h) \geq 2^{\Omega(\log^\gamma N_h)} ,$$

kde $\gamma = \log_2 3 \approx 1.58496$.

Důkaz. Z Věty 2.2 plyne

$$\mathrm{dt}(F_h) \geq 2^{\Omega(3^h)} .$$

Z předchozího výpočtu víme, že $N_h = 2 \cdot 3^{2^h - 1}$, a tedy

$$\log_2 N_h = (2^h - 1)\gamma + O(1) = \Theta(2^h) ,$$

což implikuje

$$3^h = \Theta(\log_2^\gamma N_h) .$$

Dosazením do odhadu z Věty 2.2 dostaneme dolní odhad $\mathrm{dt}(F_h)$ v požadovaném tvaru. □

Až vypočteme $\mathrm{dnf}(G_h)$ a $\mathrm{dnf}(\neg G_h)$, pak dostaneme podobným způsobem pro funkci G_h následující silnější odhad.

Věta 3.2 *Pro libovolné h nechť $N_h = \mathrm{dnf}(G_h) + \mathrm{dnf}(\neg G_h)$. Pak platí*

$$\mathrm{dt}(G_h) \geq 2^{\Omega(\log^2 N_h)} .$$

Připomeňme, že jsme již dříve pro libovolnou Booleovskou funkci f dokázali horní odhad

$$\mathrm{dt}(f) \leq 2^{O(\log^2 N \, \log n)} ,$$

kde $N = \mathrm{dnf}(f) + \mathrm{dnf}(\neg f)$ a n je počet proměnných f. Věta 3.2 dokazuje, že tento horní odhad je pro $f = G_h$ přesný až na faktor $\Theta(\log n)$ v exponentu, přičemž n je v tomto případě zhruba logaritmické vůči N.

4 Evasive funkce

Funkce f n proměnných se nazývá evasive, pokud každý rozhodovací strom pro f má hloubku n. Jinak řečeno, pro každý strom pro f existuje vstup, pro který jsou testovány všechny proměnné. Příklady takových funkcí jsou konjunkce, disjunkce nebo funkce většiny. Obecněji, každá nekonstantní symetrická funkce je evasive. Jinou postačující podmínku dává následující tvrzení.

Věta 4.1 *Nechť f je Booleovská funkce a $par(x)$ je funkce parity na stejné množině proměnných. Jestliže*

$$\sum_{par(x)=0} f(x) \neq \sum_{par(x)=1} f(x) \ ,$$

pak f je evasive.

Důkaz. Pokud existuje strom hloubky nejvýše $n-1$, který funkci vyjadřuje, pak každý jeho list přispívá stejnou hodnotou do obou sum v podmínce ve znění věty. Pokud tedy jsou tyto sumy různé, strom hloubky nejvýše $n-1$ pro f neexistuje. □

Je zřejmé, že podmínku z Věty 4.1 splňuje každá funkce, která má lichý počet jedniček ve své tabulce. Jestliže f je funkce n proměnných a f' její reprezentace pomocí $\{1, -1\}$, pak je podmínka z Věty 4.1 ekvivalentí tomu, že $\langle f', Z_{\{1,\ldots,n\}} \rangle \neq 0$. Tato podmínka znamená, že polynom, který funkci vyjadřuje, má nenulový koeficient u součinu všech proměnných a tedy má stupeň n. V této přeformulaci lze Větu 4.1 také jednoduše dokázat, protože každou funkci lze vyjádřit polynomem stupně nejvýše rovným minimální hloubce stromu, který funkci reprezentuje.

Všechny funkce, které jsou vyjádřitelné formulí v bázi $\{\wedge, \vee\}$, ve které se každá proměnná vyskytuje nejvýše jednou, jsou evasive. Důkaz lze provést tak, že v každém uzlu stromu budeme předpokládat tu hodnotu testované proměnné, která ve formuli pouze odstraní zafixovanou proměnnou, ale nevede na další zjednodušení. Lze ukázat, že tyto funkce mají lichý počet jedniček v tabulce, a tedy splňují podmínku z Věty 4.1. Příkladem symetrické evasive funkce, která nesplňuje podmínku z Věty 4.1 je $x_1x_2x_3 \vee \overline{x}_1\overline{x}_2\,\overline{x}_3$.