TIN 085 - Vybrané kapitoly z výpočetní složitosti I **zima 2006/2007**

1. domácí úlohy - opakování

Vzorové řešení

3.listopadu 2006

1. Bez použití Černovovy nerovnosti ukažte, jak se dá snížit chyba algoritmu z $1/3$ na 2^{-n}. (Hint: Snižte chybu nejprve na $1/4$ a pak použijte standardní kalkulaci.)

   Řešení: Z původního algoritmu A s chybou $1/3$ vytvoříme dva algoritmy A' a A'' pro ten samý jazyk, přičemž A' bude pracovat s chybou $< 2/9$ a A'' bude pracovat s chybou 2^{-n}. A' pracuje tak, že spustí původní algoritmus A pětkrát nezávisle a odpoví převažující odpověď. A'' pak pracuje tak, že spustí algoritmus A' $2cn+1$-krát nezávisle, pro vhodnou konstantu $c > 1$, a odpoví převažující odpověď.

   Algoritmus A' udělá chybu s pravděpodobností:

$$\binom{5}{0}\left(\frac{1}{3}\right)^5 + \binom{5}{1}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right) + \binom{5}{2}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2 = \left(\frac{1}{3}\right)^5(1+5\cdot 2+10\cdot 4) < \frac{2}{9}.$$

   Algoritmus A'' udělá chybu s pravděpodobností menší než $\sum_{i=cn+1}^{2cn+1}\binom{2cn+1}{i}\left(\frac{2}{9}\right)^i < 2^{2cn+1}\left(\frac{2}{9}\right)^{cn+1} < \left(\frac{8}{9}\right)^{cn+1}$. Pro $c > -\log_{\frac{8}{9}} 2$ je tedy chyba algoritmu A'' pod 2^{-n}.

2. Nechť M je náhodná matice vzatá z uniformního rozdělení matic $n \times n$ nad tělesem $GF[2]$. Jaká je pravděpodobnost, že M je regulární (invertibilní)? Uveďte přesnou formuli a rozumné odhady na tuto pravděpodobnost pro $n \to \infty$. Udělejte to samé pro matice nad $GF[p]$. Co se děje s danou pravděpodobností, když $p \to \infty$?

   Řešení: k lineárně nezávislých vektorů z $GF[2]^n$ generuje podprostor $GF[2]^n$ velikosti 2^k. Pro každé $k \in \{1,\dots,n\}$, prvních k řádků regulární matice M musí být lineárně nezávislé vektory z $GF[2]^n$. První řádek M může být tedy libovolný vektor kromě nulového, druhý může být jeden z 2^n-2, třetí jeden z 2^n-2^2, obecně k-tý může být jeden z 2^n-2^{k-1}. Tedy počet regulárních matic $n\times n$ nad tělesem $GF[2]$ je $\prod_{k=1}^n(2^n-2^{k-1}) = 2^{n^2}\prod_{k=1}^n(1-2^{k-1-n})$. Celkový počet matic $n\times n$ nad tělesem $GF[2]$ je 2^{n^2}, tedy pravděpodobnost, že náhodná matice je regulární, je $\prod_{k=1}^n(1-2^{-k})$.

   Abychom tuto pravděpodobnost nějak odhadli, můžeme použít následující snadno dokazatelnou nerovnost:

   - pro $0 \le a_1, a_2, \dots, a_n \le 1$, $\prod_{k=1}^n(1-a_k) \ge 1-\sum_{k=1}^n a_k$.

   Také víme, $\sum_{k=3}^n 2^{-k} \le \sum_{k=3}^\infty 2^{-k} = \frac{1}{4}$. Tudíž $\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{3}{4} \le \prod_{k=1}^n(1-2^{-k}) \le \frac{1}{2}\cdot\frac{3}{4}\cdot\frac{7}{8}$, to jest pravděpodobnost, že náhodná matice velikosti $n\times n$ nad tělesem $GF[2]$ je regulární je mezi 0.28125 a 0.328125.

   Podobně pro matice nad $GF[p]$ dostaneme, že je zde $p^{n^2}\prod_{k=1}^n(1-p^{k-1-n})$ regulárních matic. Tedy pro pravděpodobnost q, že náhodně vybereme regulární $n\times n$ matici nad $GF[p]$, platí:

$$(1-\frac{1}{p})\cdot(1-\frac{1}{p^2})\cdot\left(1-\frac{1}{p^3}\cdot\frac{1}{1-\frac{1}{p}}\right) \le q = \prod_{k=1}^n(1-2^{-k}) \le (1-\frac{1}{p})\cdot(1-\frac{1}{p^2}).$$

   Když $p\to\infty$, pravděpodobnost $q\to 1$, to jest náhodná matice bude téměř jistě regulární.

3. Vezměme si celá čísla $0 < r < n$ a vyberme uniformě náhodně r vektorů $u_1, u_{2,}, \dots, u_r$ z $GF[2]^n = \{0,1\}^n$.

   (a) Jaká je pravděpodobnost, že $u_1,\dots,u_r$ jsou lineárně nezávislé?

   (b) Označme písmenem B_ϵ množinu nenulových vektorů, jež jsou lineární kombinací $u_1,\dots,u_r$ a jež mají Hammingovu váhu méně než ϵn, kde $0 < \epsilon < 1/2$ je zvolená konstanta. Jaká je očekávaná velikost množiny B_ϵ? (Hammingova váha je počet jedniček ve vektoru.)

(Povšimněte si, že pokud všechny nenulové lineární kombinace vektorů u_i mají Hammingovu váhu větší nebo rovnu ϵn, tj. $B_\epsilon = \emptyset$, pak tyto vektory (lineární kombinace) tvoří *samoopravný kód* se vzdáleností ϵn, to jest kód schopný opravit $(\epsilon n - 1)/2$ chyb. Takový kód existuje pokud $r < (1 - H(\epsilon))n$, kde $H(\epsilon) = -\epsilon \log \epsilon - (1 - \epsilon) \log(1 - \epsilon)$. Rozmyslete si proč.)

Řešení:

(a) Z předchozího cvičení víme, že $u_1, \ldots, u_r$ jsou lineárně nezávislé s pravděpodobností $\prod_{k=1}^{r}(1 - 2^{k-1-n}) = \prod_{k=n+1-r}^{n}(1 - 2^{-k})$. Například pro $r < n - \log n$ je tato pravděpodobnost vetší než $(1 - \frac{2}{n})$ a pro r ještě menší dále roste zhruba jako $(1 - 2^{r-n+1})$.

(b) Jako obvykle nás až tak nezajímá přesný výraz jako spíše rozumný odhad. Vzhledem k nastíněné aplikaci nás vskutku zajímá horní mez na očekávanou velikost množiny B_ϵ . Nejprve potřebujeme odhadnout velikost množiny $H_{\epsilon n} = \{x \in \{0,1\}^n; x$ má Hammingovu váhu nejvýše $\epsilon n\}$. To se dá udělat s užitím Černovovi nerovnosti, kdy si za náhodné proměnné zvolíme

$$X_i = \begin{cases} 1 & \text{s pravděp. } 1/2 \\ 0 & \text{s pravděp. } 1/2 \end{cases}$$

a ptáme se, jaká je pravděpodobnost, že $\sum X_i \leq \epsilon n$. S užitím verze Černovovi nerovnosti dané na přednášce, $Pr[\sum X_i \leq \epsilon n] \leq e^{-(1-\epsilon)^2 n/2}$. To jest

$$|H_{\epsilon n}| \leq 2^n e^{-(1-\epsilon)^2 n/2} = 2^{(1-\frac{1}{2}(1-\epsilon)^2 \log_2 e)n}.$$

Vezměme si nějakou lineární kombinaci $v = \sum a_i u_i$, pro pevná a_i. Když vektory u_i jsou vybrány náhodně, pak v je také náhodný vektor distribuovaný uniformě v $GF[2]^n$. Pravděpodobnost, že v padne do $H_{\epsilon n}$ je tedy nejvýše $e^{-(1-\epsilon)^2 n/2}$. To platí pro každou lineární kombinaci vektorů u_i. Pro dané $a_1, \ldots, a_r \in \{0,1\}$, můžeme definovat náhodnou proměnnou:

$$X_{a_1,\ldots,a_r} = \begin{cases} 1 & \text{když } \sum a_i u_i \in H_{\epsilon n}, \\ 0 & \text{jinak.} \end{cases}$$

Zřejmě $|B_\epsilon| = \sum X_{a_1,\ldots,a_r}$. Snadno nahlédneme, že $E_{u_1,\ldots,u_r}[X_{a_1,\ldots,a_r}] \leq e^{-(1-\epsilon)^2 n/2}$. Tudíž $E[|B_\epsilon|] \leq 2^r e^{-(1-\epsilon)^2 n/2}$. To je to, co jsme chtěli vědět.

Pokud $r < \frac{n}{2}(1-\epsilon)^2 \log_2 e$, pak $E[|B_\epsilon|] < 1$ a tudíž pro nějaký výběr vektorů $u_1, \ldots, u_r$ je B_ϵ prázdné. V takovém případě v Hammingově blízkosti žádné z lineárních kombinací vektorů u_i neleží žádná jiná lineární kombinace u_i, neboť jejich rozdíl by byla též jedna z lineárních kombinací a musela by mít malou Hammingovu váhu. Pokud vezmeme r ještě o jedna menší, bude $E[|B_\epsilon|] < 1/2$ a podle Markovovi nerovnosti tak B_ϵ bude prázdné pro více než polovinu náhodných výběrů vektorů $u_1, \ldots, u_r$. Zároveň pak můžeme požadovat, aby $u_1, \ldots, u_r$ byly lineárně nezávislé. Takto lze sestrojit *lineární samoopravné kódy* se schopností opravit konstantní zlomek chyb a s konstantní kapacitou.

Lepší odhad na velikost $H_{\epsilon n}$ je následující:

$$\frac{2^{H(\epsilon)n}}{n+1} \leq |H_{\epsilon n}| \leq 2^{H(\epsilon)n}.$$

Zde $H(\epsilon) = -\epsilon \log \epsilon - (1-\epsilon) \log(1-\epsilon)$ je takzvaná *binární entropie*. Tento vztah se dá relativně snadno dokázat. Předně:

$$|H_{\epsilon n}| = \sum_{i=0}^{\epsilon n} \binom{n}{i}.$$

Dále pak pro $k = \epsilon n$ podle Binomické věty:

$$1 = 1^n = \left(\frac{n-k}{n} + \frac{k}{n}\right)^n = \sum_{i=0}^{n} \binom{n}{i} \left(\frac{n-k}{n}\right)^i \left(\frac{k}{n}\right)^{n-i}.$$

Přezkoumáním poměru po sobě jdoucích členů v sumě zjistíme, že největší člen je pro $i = n-k$. Tedy $1 \leq (n+1)\binom{n}{n-k}(\frac{n-k}{n})^{n-k} \left(\frac{k}{n}\right)^k = (n+1)\binom{n}{k}2^{H(\epsilon)n}$. Z toho vyplývá dolní odhad na velikost $H_{\epsilon n}$. Horní odhad se dostane obdobně.

4. Cílem tohoto cvičení je zopakovat, jak se konstruují konečná (komutativní) tělesa. Připomeňme si *Euklidův algoritmus*, pro hledání největšího společného dělitele dvou celých čísel a a b. Vstupem algoritmu je $a > b > 0$. Algoritmus definuje posloupnost $a_0, a_1, \ldots$ takto:

- $a_0 = a$, $a_1 = b$ a
- pro $i = 2, 3, \ldots$ nechť $a_i = a_{i-2} \bmod a_{i-1}$.

Nechť t je nejmenší takové, že $a_t = 0$. Výstupem algoritmu je a_{t-1}.

(a) Dokažte korektnost Euklidova algoritmu, to jest, že t je dobře definované a že výstupem je největší společný dělitel a a b.

(b) Ukažte, že když $nsd(a, b) = c$, pak existují celá čísla x a y takové, že $|x| \leq b$, $|y| \leq a$ a $ax + by = c$.

(c) Ukažte, že pro prvočíslo p, $GF[p] = \{0, 1, \ldots, p-1\}$ s počítáním modulo p je těleso. (Tedy především ukažte, že $GF[p]^* = GF[p] \setminus \{0\}$ tvoří grupu při násobení modulo p.)

(d) Nechť F je konečné těleso a $F_n[x] = \{a_0 + a_1 x + \cdots + a_{n-1} x^{n-1}; a_0, a_1, \ldots, a_{n-1} \in F\}$ jsou polynomy nad F stupně nejvýše $n-1$. Nechť $r(x)$ je ireducibilní polynom stupně n. Ukažte, že $F_n[x]$ s počítáním modulo $r(x)$ je těleso. (Hint: Ukažte, že Euklidův algoritmus funguje i pro polynomy. Polynom je ireducibilní, když nelze zapsat jako součin jiných dvou.)

Řešení:

(a) Vzhledem k tomu, že zbytek při dělení je vždy nezáporný a menší než dělitel, indukcí snadno nahlédneme, že posloupnost $a_0, a_1, a_2, \ldots$ je nezáporná a striktně klesající, dokud se nenarazí na $a_t = 0$. Tudíž t je dobře definované. Dále pak se snadno ověří, že $nsd(a_i, a_{i+1}) = nsd(a_{i+1}, a_{i+2})$, pokud $a_{i+1} \neq 0$. Toto ověření je nutné udělat, mi ho však zde přeskočíme. Z toho vyplývá, že $nsd(a, b) = nsd(a_0, a_1) = nsd(a_{t-1}, a_t) = a_{t-1}$.

(b) Opět lze indukcí snadno ověřit, že pro každé $i = 0, 1, \ldots$, existuje x_i a y_i takové, že $a_i = a_0 x_i + a_1 y_i$. To vyplývá z toho, že a_i lze získat z a_{i-2} postupným odečítáním a_{i-1}. Mez na velikost potřebných x_i a y_i vyplývá z toho, že pokud $y_i > a_o$ pak můžeme x_i zvětšit o a_1 a y_i zmenšit o a_0, při zachování požadovaných vztahů. Obdobně pro $-y_i > a_o$. To lze dělat tak dlouho, dokud nenastane $|y_i| < a_o$. Jelikož a_i je menší než a_1, x_i musí být také malé.

(c) Pro p prvočíslo, všechna čísla m mezi 1 a $p-1$ jsou nesoudělná s p. Tudíž $nsd(m, p) = 1$. Podle předchozí části pak existuje m' takové, že $mm' = 1 (\bmod p)$. Tedy ke každému nenulovému prvku z $GF[p]$ lze nalézt inverzní prvek. Díky asociativitě, komutativitě a $1 \cdot 1 = 1$, $GF[p]^*$ musí být grupa. Detaily zde přeskočíme.

(d) Postupuje se podle hintu a ukáže se (a), (b) a (c) také pro polynomy, přičemž se za velikost polynomů považuje jejich stupeň.