

POSUDEK NA DIPLOMOVOU PRÁCI MARTINA
BABKY "PROPERTIES OF UNIVERSAL HASHING"

Předložená diplomová práce se zabývá problémem délky nejdelšího řetězce v univerzálním hašování. Význam tohoto problému dává Markovova nerovnost, která by v této situaci umožňovala, abychom při několikanásobném překročení očekávané délky maximálního řetězce opakovali volbu hašovací funkce. Předloženou práci můžeme rozdělit do tří částí. V první části jsou uvedeny základní definice a fakta o hašování, přičemž speciální pozornost je věnována univerzálnímu hašování. Jsou zde např. definovány různé univerzální systémy. Tato část je rozdělena do čtyř kapitol. Vlastní jádro práce tvoří pátá kapitola, která vychází z výsledku Alonga, Dietzfelbingera, Bro Miltersena, Petranka a Tardose. Tento výsledek říká, že když použijeme jako univerzální systém lineární transformace mezi vektorovými prostory nad tělesem \mathbb{Z}_2 a když v tabulce velikosti m reprezentujeme množinu velikosti nejvýše $m \log m$, pak očekávaná délka nejdelšího řetězce je $O(\log m \log \log m)$. Tento výsledek je jen teoretický, protože spočítané konstanty jsou příliš velké. Autor tento výsledek použil a vylepšil tyto konstanty tak, že je prakticky použitelný. Navíc přepsal původní důkaz do čitelnější podoby. V poslední části práce (šestá a sedmá kapitola) autor diskutuje praktickou použitelnost uvedených výsledků. Diskutuje použitelnost postupu zmíněného na začátku tohoto posudku a na závěr ukazuje na možná pokračování této práce. Práce navíc obsahuje dva dodatky, jeden z nich připomíná základní definice a fakta z lineární algebry a druhý se věnuje teorii pravděpodobnosti.

Práce je napsaná anglicky. Jazyk je trochu kostrbatý a místy jsou vidět doslovné překlady (protože je to první odborná práce autora v anglickém jazyce, tak je to pochopitelné a vlastně velmi dobré). Dále uvedu některé konkrétní nedostatky:

V páté kapitole autor soustavně odkazuje na [24] místo správného odkazu [3].

Je dost nezvyklé psát Claim 5.12, Claim 5.11 místo obvyklého Claims 5.11 and 5.12 (str. 44, řádek 4 zdola).

Angličtina na str. 45, řádek 9 zdola, je nesrozumitelná.

Na str. 49 na řádcích 18 a 17 zdola má být \mathbb{Z}_2^u místo U a \mathbb{Z}_2^f místo F .

Na str. 51 na řádce 6 zdola má být $\binom{|S|}{2} 2^{-f}$ místo $\binom{|S|}{2} 2^{-u}$.

Na str. 63, řádek 11, má být Corollary 5.34 místo Remark 5.36.

Na str. 75 na řádce 10 myslím, že vypadlo *at most*.

Na str. 77 na řádce 12 má být *superset* místo *subset*.

Na str. 90, řádek 11, má být *The next α -cycle*, takhle je to hodně nesrozumitelné. Dále na této stránce p znamená potenciál, ale před tím se tento symbol používal k označení pravděpodobnosti.

Na str. 99 v Definici A3 jsou popletené symboly U a V (asi na začátku mají být prohozené).

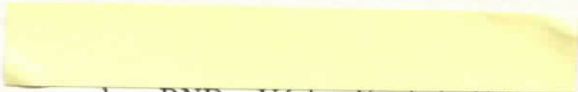
Na str. 100 Lemma A.10 je zbytečně omezená. Není třeba, aby T bylo na, obecně platí $|T^{-1}(y)| = 2^{f-a}$, kde a je dimenze obrazu T .

Na str. 101, 1. řádek, má být $T^{-1}(0)$ nikoliv $T^{-1}(0)$.

Na str. 106, řádek 4, mělo být řečeno, že f je probability density function.

Obecně bych řekl, že autor má v oblibě používat Remark místo Lemma nebo Proposition, takže Remark velmi často označuje důležité tvrzení místo poznámky. V Claim 6.15 chybí tvrzení: If $p_0 = 0$ and $p \geq 0$ then $A_o = T_o$. Záporné otázky (běžné v našich jazycích) přeložené do angličtiny jsou nesrozumitelné. Největší problém mám na str. 76, protože nechápu, proč lze vynechat $\log \alpha$.

Shrnutí: práce vysoko překračuje úroveň diplomových prací, přináší zajímavé a netriviální zobecnění hlubokého výsledku. Proto jednoznačně doporučuji ji **uznat jako diplomovou práci**.



doc. RNDr. Václav Koubek, DrSc.