Varianty pojmu užitočnosti informácie

Vincent Hlaváč

Školiteľ: prof. RNDr. Branislav Rovan PhD.

Podstata pojmu

Informáciu považujeme za užitočnú, ak nám pomôže riešiť problém ľahšie.

Zdroj: Giovanni Pighizzini, Branislav Rovan and Šimon Sádovský. Usefulness of information and decomposability of unary regular languages. Information and Computation, page 104868, 2022

Formalizácia a stav problematiky

Formalizácia

- problém - jazyk L, respektíve zistiť či $w \in L$
- riešenie - automat A akceptujúci L ($L(A) = L$)
- zložitosť riešenia - stavová zložitosť automatu A ($sc(A)$)
- rada - jazyk L_{adv} akceptovaný automatom A_{adv} ($L(A_{adv}) = L_{adv}$)
- nový problém - jazyk L_{new} akceptovaný automatom A_{new} ($L(A_{new}) = L_{new}$)
- $L_{adv} \cap L_{new} = L$
- $sc(A_{new}) < sc(A_{min})$, $sc(A_{adv}) < sc(A_{min})$

Formalizácia

- problém - jazyk L, respektíve zistiť či $w \in L$
- riešenie - automat A akceptujúci L ($L(A) = L$)
- zložitosť riešenia - stavová zložitosť automatu A ($sc(A)$)
- rada - jazyk L_{adv} akceptovaný automatom A_{adv} ($L(A_{adv}) = L_{adv}$)
- nový problém - jazyk L_{new} akceptovaný automatom A_{new} ($L(A_{new}) = L_{new}$)
- $L_{adv} \cap L_{new} = L$
- $sc(A_{new}) < sc(A_{min})$, $sc(A_{adv}) < sc(A_{min})$

Budeme skúmať prípad DFA.

Definícia

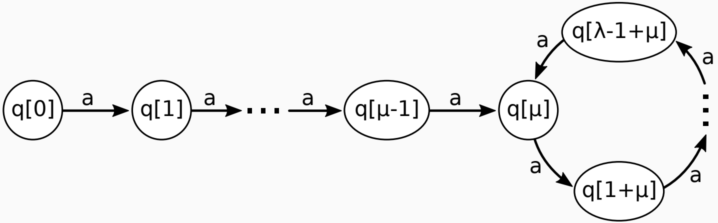
Nech A je DFA. Dvojicu DFA $[A_1, A_2]$ nazývame *2-rozklad* DFA A ak $L(A) = L(A_1) \cap L(A_2)$, $sc(A_1) < sc(A)$ a $sc(A_2) < sc(A)$. Ak takýto 2-rozklad A existuje, hovoríme, že A je *2-rozložiteľný*.

Definícia

Nech $L \in \mathcal{R}$. Hovoríme, že L je deterministicky 2-rozložiteľný ak minimálny DFA akceptujúci L je 2-rozložiteľný.

Unárne DFA

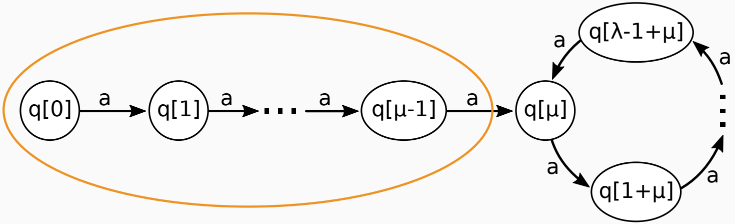
◇ UDFA - DFA s unárnou vstupnou abecedou



Obr. 1: DFA A veľkosti (λ, μ)

Unárne DFA

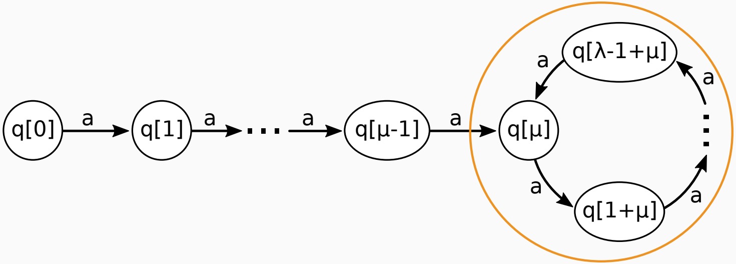
- UDFA sa skladá z *chvosta*



Obr. 2: DFA A veľkosti (λ, μ)

Unárne DFA

- UDFA sa skladá z chvosta a *cyklu*



Obr. 3: DFA A veľkosti (λ, μ)

Unárne DFA

- veľkosť UDFA je dvojica (λ, μ), kde
  - λ je počet stavov cyklu
  - μ je počet stavov chvosta
- minimálne UDFA - minimálna veľkosť cyklu aj chvosta

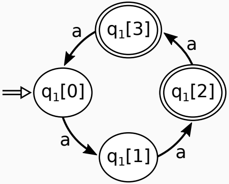
Unárne DFA

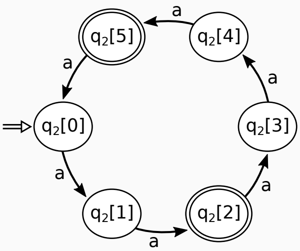
Definícia

Nech L je unárny regulárny jazyk, $\lambda \in \mathbb{N}$. Hovoríme, že L je λ-*cyklický* ak existuje UDFA A veľkosti $(\lambda, 0)$ taký, že $L(A) = L$. L nazývame *prísne* λ-*cyklický* ak je λ-cyklický a neexistuje $\lambda' < \lambda$ také, že L je λ'-cyklický.

2-rozložiteľnosť - príklad

- $L = \{a^{12k+2}, a^{12k+11} \mid k \in \mathbb{N}\}$
- minimálny automat akceptujúci L je UDFA A veľkosti $(12, 0)$
- L je prísne λ-cyklický jazyk pre $\lambda = 12$





Kvalita rozkladov

Kvalita rozkladov

- označme $\mathcal{D}$ množinu n-tíc automatov takých, že tvoria n-rozklad nejakého DFA A.
- funkciu $f : \mathcal{D} \to \mathbb{N}$ nazveme *f-kvalita rozkladu*

Definícia

Nech $D = [A_1, \ldots, A_n]$, $D \in \mathcal{D}$ je n-rozklad nejakého DFA A. Funkciu $sum : \mathcal{D} \to \mathbb{N}$, definovanú ako

$$sum(D) = \sum_{i=1}^{n} sc(A_n)$$

nazývame *sum*-kvalita rozkladu.

Efektívne 2-rozklady

- rozklady určitého tvaru pre prísne λ-cyklické jazyky
- UDFA A veľkosti $(\lambda, 0)$
- UDFA A_1, A_2 veľkostí $(\lambda_1, 0), (\lambda_2, 0)$ kde $lcm(\lambda_1, \lambda_2) = \lambda$

Výsledky

- Každý 2-rozklad, kde A_1, A_2 sú veľkostí $(\lambda_1, \mu_1), (\lambda_2, \mu_2)$ sa dá upraviť na efektívny 2-rozklad, kde B_1, B_2 sú veľkostí $(\lambda_1/\alpha_1, 0), (\lambda_2/\alpha_2, 0)$ pre nejaké $\alpha_1, \alpha_2 \in \mathbb{N}$.
- Pre každý 2-rozklad $[A_1, A_2]$ existuje efektívny 2-rozklad $[B_1, B_2]$ taký, že $\mathit{sum}([A_1, A_2]) \geq \mathit{sum}([B_1, B_2])$.

Sum-kvalita 2-rozkladov

Veta

Nech $[A_1, A_2]$ *je efektívny 2-rozklad minimálneho automatu* A *pre nejaký prísne* λ*-cyklický jazyk* L *pre nejaké* $\lambda \in \mathbb{N}$*. Potom* $\lceil 2\sqrt{\lambda} \rceil \leq sum([A_1, A_2]) \leq \frac{5}{6}\lambda$.

Idea dôkazu $sum([A_1, A_2]) \leq \frac{5}{6}\lambda$

- $sc(A_1) = \lambda_1,\ sc(A_2) = \lambda_2$
- $lcm(\lambda_1, \lambda_2) = \lambda,\ \lambda_1 < \lambda,\ \lambda_2 < \lambda$
- λ_1 / λ teda $\lambda = \alpha\lambda_1$ a tiež $\alpha > 1$
- λ_2 / λ teda $\lambda = \beta\lambda_2$ a tiež $\beta > 1$
- $gcd(\alpha, \beta) = 1$

Idea dôkazu $sum([A_1, A_2]) \leq \frac{5}{6}\lambda$

- $sc(A_1) = \lambda_1$, $sc(A_2) = \lambda_2$
- $lcm(\lambda_1, \lambda_2) = \lambda$, $\lambda_1 < \lambda$, $\lambda_2 < \lambda$
- λ_1/λ teda $\lambda = \alpha\lambda_1$ a tiež $\alpha > 1$
- λ_2/λ teda $\lambda = \beta\lambda_2$ a tiež $\beta > 1$
- $gcd(\alpha, \beta) = 1$

Idea dôkazu $sum([A_1, A_2]) \leq \frac{5}{6}\lambda$

- $sc(A_1) = \lambda_1$, $sc(A_2) = \lambda_2$
- $lcm(\lambda_1, \lambda_2) = \lambda$, $\lambda_1 < \lambda$, $\lambda_2 < \lambda$
- λ_1/λ teda $\lambda = \alpha\lambda_1$ a tiež $\alpha > 1$
- λ_2/λ teda $\lambda = \beta\lambda_2$ a tiež $\beta > 1$
- $gcd(\alpha, \beta) = 1$

Idea dôkazu $sum([A_1, A_2]) \leq \frac{5}{6}\lambda$

λ_1, λ_2 sú najväčšie vtedy, keď α, β najmenšie, teda $2, 3$

$$2\lambda_1 \leq \lambda \quad 3\lambda_2 \leq \lambda$$

$$\lambda_1 \leq \frac{1}{2}\lambda \quad \lambda_2 \leq \frac{1}{3}\lambda$$

$$\lambda_1 + \lambda_2 \leq \frac{1}{2}\lambda + \frac{1}{3}\lambda = \left(\frac{1}{2} + \frac{1}{3}\right)\lambda = \frac{5}{6}\lambda$$

Výsledky

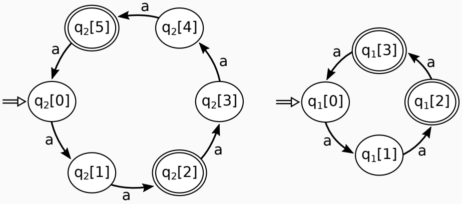
- Sum-kvalita najlepšieho 2-rozkladu môže byť od dolného odhadu ľubovoľne vzdialená.
- Horné a dolné odhady sum-kvality efektívnych 2-rozkladov sú presné.

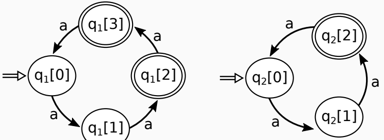
Presnosť odhadov

$$L = \{a^{12k+2}, a^{12k+11} \mid k \in \mathbb{N}\}$$

$$\lambda_1^{A_1} + \lambda_2^{A_2} = 4+6 = 10 = \frac{5}{6}12 = \frac{5}{6}\lambda$$

$$\lambda_1^{B_1} + \lambda_2^{B_2} = 4+3 = 7 = \lceil 2\sqrt{12} \rceil = \lceil 2\sqrt{\lambda} \rceil$$





3-rozložiteľnosť

3-rozložiteľnosť a efektívne 3-rozklady

- 3-rozložiteľnosť definovaná obdobne ako 2-rozložiteľnosť
- navyše: $L(A_i) \cap L(A_j) \neq L(A)$ pre $i \neq j$.
- efektívne 3-rozklady definované obdobne ako efektívne 2-rozklady
- navyše: $lcm(\lambda_i, \lambda_j) < \lambda$ pre $i \neq j$
- nie všetky 3-rozložiteľné jazyky sú aj efektívne 3-rozložiteľné

3-rozložiteľnosť a efektívne 3-rozklady

- 3-rozložiteľnosť definovaná obdobne ako 2-rozložiteľnosť
- navyše: $L(A_i) \cap L(A_j) \neq L(A)$ pre $i \neq j$.
- efektívne 3-rozklady definované obdobne ako efektívne 2-rozklady
- navyše: $lcm(\lambda_i, \lambda_j) < \lambda$ pre $i \neq j$
- nie všetky 3-rozložiteľné jazyky sú aj efektívne 3-rozložiteľné

Výsledky

- Charakterizácia efektívne 3-rozložiteľných jazykov.
- Ak je jazyk efektívne 3-rozložiteľný, tak je aj efektívne 2-rozložiteľný, nie však opačne.

Porovnanie kvality rozkladov

Výsledky

- $\lceil 3\sqrt[3]{\lambda} \rceil \leq sum([A_1, A_2, A_3]) \leq \frac{1}{3}\lambda$
- Horné a dolné odhady sum-kvality efektívnych 3-rozkladov sú presné.
- Existuje jazyk taký, že každý jeho efektívny 3-rozklad je kvalitnejší ako každý jeho jeho efektívny 2-rozklad.
- Existuje jazyk taký, že existuje jeho efektívny 3-rozklad, ktorý je kvalitnejší ako každý jeho 2-rozklad, a tiež existuje jeho efektívny 3-rozklad a existuje jeho efektívny 2-rozklad také, že daný 2-rozklad je kvalitnejší ako daný 3-rozklad.

Výsledky

- $\lceil 3\sqrt[3]{\lambda} \rceil \leq sum([A_1, A_2, A_3]) \leq \frac{1}{3}\lambda$
- Horné a dolné odhady sum-kvality efektívnych 3-rozkladov sú presné.
- Existuje jazyk taký, že každý jeho efektívny 3-rozklad je kvalitnejší ako každý jeho jeho efektívny 2-rozklad.
- Existuje jazyk taký, že existuje jeho efektívny 3-rozklad, ktorý je kvalitnejší ako každý jeho 2-rozklad, a tiež existuje jeho efektívny 3-rozklad a existuje jeho efektívny 2-rozklad také, že daný 2-rozklad je kvalitnejší ako daný 3-rozklad.

Ďakujem

Efektívne 2-rozložiteľné jazyky, ktoré nie sú efektívne 3-rozložiteľné

- Problém spočíva v malom počte rôznych prvočíselných faktorov λ, konkrétne daná situácia nastáva pri dvoch.
- $L^{i,j}_{p,q} = \{a^{p^i q^j k} \mid k \in \mathbb{N}\}$
- $\{L^{i,j}_{p,q} \mid p, q \textit{ sú prvočísla};\ i, j \in \mathbb{N};\ i, j \geq 1\}$
- Môžu to byť aj zložitejšie jazyky s viac ako jedným akceptačným stavom v minimálnom automate