

Toteutusdokumentti

Toteutus ja analysointi

Ohjelman perusideana on, että se osaa luoda suomen- ja englanninkielien kuuloisia nimiä. Kun halutaan luoda nimi Nimeaja- luokan avulla ilmoitetaan, kummalla kielellä nimi halutaan, minkä pituisia nimiä halutaan ja kuinka pirkä markovin ketju luodaan. Tämän jälkeen Nimeäjä lukee Opetus- lukan luoman tiedoston, jonne on luotu Trie-rakenteena puu, jossa on kaikki opetusdatasta otetut merkkijonorimpsut. Sitten Nimeaja luo sen pohjalta uuden nimen, käyttäen todennäköisyyksiä seuraavan kirjaimen poiminnassa (eli markovin ketjujen avulla). Näitä nimiä luodaan 20.

Aikavaativuudeltaan algoritmia voidaan tutkia kahdelta osalta: Puun luominen ja nimen luominen. Puun luomisen aikavaativuus on riippuvainen siitä, kuinka monta nimeä käydään läpi puun luonnin aikana. Se on myös riippuvainen nimen pituudesta, mutta koska nimet ovat yleisesti 3- 15 kirjaimisia, nimestä tulee maksimissaan noin 12 eri merkkijonoa, joka ei vaikuta aikavaativuuteen kovinkaan paljon. Aikavaativuus on $O(n)$, koska opetuksessa käydään raa’asti jokainen sana läpi, eikä algoritmissa ole esimerkiksi rekursiota missään kohtaa. Nimen luominen on nopeampaa. Se on riippuvainen siitä, minkä pituinen nimi halutaan. For-luuppeja käydään läpi lasta etsiessä, mutta koska lapsia voi olla vain yhtä monta, kuin kirjaimia jossain kielessä, lapsien maksimimäärä on tiedossa. Tällöin yhden nimen luontiin kestää pituus*puun läpikäynti lehtisolmuun asti + vakioaikaisia toimintoja, eli prosessi on vakioaikainen $O(n*k)$, jossa voidaan olettaa, että $n < 15$ ja k on kirjaimia kyseisessä kielessä

Työssä on se puute, että kun kielen nimet ovat kovin erikuuluisia (kuten englanninkielessä) tuloksista tulee epätarkempia ja kuulostavat vähemmän alkuperäisen kielen nimiltä. Tällöin vain pitää valita pitempi ketju. Työtä voisi parannella luomalla tiettyjä kielioppisääntöjä jokaiselle kielelle (esim ei kahta samaa konsonanttia nimen alkuun tai loppuun suomen kielessä). Myös käyttäjän toiveita voisi kuunnella enemmän, esim alkukirjaimessa ja nimen sukupuolella. Myös hienomman grafiikan voisi luoda, missä painellaan vain nappuloita ja valitaan kieli tai pituus, mutta siihen ei aika riittänyt.

Lähteet

Markov chains:

<https://www.youtube.com/watch?v=i3AkTO9HLXo>

<https://www.youtube.com/watch?v=VNHeFp6zXKU>

Nimet:

https://lapsennimi.com/naisten_nimet_tyton_nimia_lista/

https://lapsennimi.com/miesten_nimet_pojan_nimia_lista/

<https://www.verywellfamily.com/top-1000-baby-boy-names-2757618>

<https://www.goodhousekeeping.com/life/parenting/a37668901/top-baby-girl-names/>