

Toteutusdokumentti

Miten ohjelma on toteutettu?

Tässä ohjelmassa säännöllisten lauseiden tulkkaus perustuu epädeterministisen säännöllisen automaatin rakentamiseen. Niiden avulla voidaan selvittää vastaako annettu syöte säännöllistä lauseketta.

Ensiksi ohjelma lisää annettuun säännölliseen lausekkeeseen liitosmerkit, jotta lausekkeen muuttaminen käänteiseksi puolalaiseksi notaatioksi sujuisi helpommin. Sitten syöte, johon on lisätty liitosmerkit, annetaan metodille, joka muuttaa syötteen käänteiseksi puolalaiseksi notaatioksi. Muuntamisessa käytetään muokattua shunting-yard -algoritmia, joka osaa käsitellä säännöllisiä lausekkeita.

Käänteisen säännöllisen lausekkeen perusteella luodaan automaatti. Käänteisen notaation etu on se, että käänteisen lausekkeen lukemisessa ei tarvita rekursiota ja toteutus on paljon selkeämpi. Automaatin rakennus perustuu Ken Thompsonin vuonna 1968 kehittämään algoritmiin.

Kun automaatti on valmis, voidaan sille antaa syötteitä. Automaatin simuloinnissa käytetään omaa algoritmiani, joka vaikuttaa onneksi toimivan kaikilla syötteillä. Lopuksi automaatti palauttaa true tai false sen mukaan, hyväksyikö se syötteen vai ei.

Työn puutteet

Ohjelma ei tällä hetkellä tue läheskään kaikkia ominaisuuksia, jotka on toteutettu Javan omassa versiossa. Puutteet olisi kuitenkin melko helppo korjata, mutta sitä ei lähdetty tämän harjoitustyön puitteissa tekemään. Syötteeksi ohjelmalle voi antaa myös ainoastaan kirjaimia, eikä tulkki tunnista virheellisiä säännöllisiä lausekkeita.

Lisäksi automaattia ei optimoida lainkaan. Nykyisellään automaatteihin muodostuu turhia tyhjän merkkijonon vastaanottavia siirtofunktioita. Suorituskyky vaikuttaa siitä huolimatta hyvältä.

Lähteet

<http://perl.plover.com/Regex/article.html>

http://en.wikipedia.org/wiki/Shunting_yard_algorithm

http://en.wikipedia.org/wiki/Thompson's_construction_algorithm