Heimaverkefni 5

TÖL301G Formleg mál og reiknanleiki

Hjörvar Sigurðsson

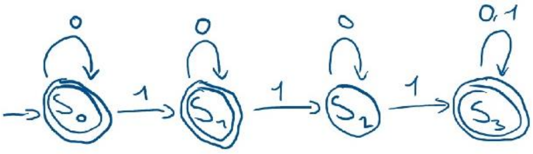
Rætt var um verkefnið við Arnar Sigurðsson

1. Já, HTML er samhengislaust mál. HTML er afleiða af SGML (Standard Generalized Markup Language), en SGML er samhengislaust mál. Það að HTML sé samhengislaust mál fæst staðfest með vitneskjunni um að hægt sé að parsa (e. parsing) HTML með því að útbúa LL málllýsingu (e. LL grammar). Þó ber að hafa í huga að einungis *gilt* HTML er samhengislaust mál – vafrar reyna oft að láta ógilt HTML duga, en í þeim tilvikum er HTML-ið oft ekki samhengislaust.

   Upplýsingar fengnar að hluta til frá https://resultfor.dev/456189-is-html-a-context-free-language.

2.

Löggeng stöðuvél M:



i. Fjögur ástönd í M, þannig ég innleiði fjórar breytur, R_0, R_1, R_2, og R_3.

ii. Þrjár stöðufærslur í M, þannig að

$$R_0 \rightarrow 0R_0 \mid 1R_1 \mid \varepsilon$$

$$R_1 \rightarrow 0R_0 \mid 1R_2 \mid \varepsilon$$

$$R_2 \rightarrow 0R_2 \mid 1R_3$$

$$R_3 \rightarrow 0R_3 \mid 1R_3 \mid \varepsilon \ .$$

iii. Þrjú samþykktarástönd í M þannig að

$$R_0 \rightarrow \varepsilon$$

$$R_1 \rightarrow \varepsilon$$

$$R_3 \rightarrow \varepsilon \ .$$

iv. S_0 upphafsástand í M, þannig að

R_0 er upphafsbreyta í G.

Niðurstaða: $G = (V, \{0,1\}, P, R_0)$

með reglur P:

$$R_0 \rightarrow 1R_1 \mid 0R_0 \mid \varepsilon$$
$$R_1 \rightarrow 1R_2 \mid 0R_1 \mid \varepsilon$$
$$R_2 \rightarrow 1R_3 \mid 0R_2$$
$$R_3 \rightarrow 1R_3 \mid 0R_3 \mid \varepsilon$$

og V: $\{R_0, R_1, R_2, R_3\}$.

3.

a.

$G = (V, \Sigma, R, S)$,

$V = \{S, A\}$

$\Sigma = \{0, 1\}$

$R = P$:

$S \rightarrow 0S11 \mid \varepsilon$

S er byrjunarstaðan. Ef að $n=0$, þá er strengurinn tómur og því ε. Ef að $n>0$, þá er hægt að velja $(0S)$ n oft en það kemur alltaf næst $2n \times (1)$.

$S = S$

Svo: $G = (\{S\}, \{0,1\}, P, S)$ með reglu $S \rightarrow 0S11 \mid \varepsilon$

b.

$G = (\{S, A, B\}, \{a, b\}, P, S)$

þar sem reglur P:

(i) $S \rightarrow aA \mid Bb$

(ii) $A \rightarrow \varepsilon \mid aA \mid aAb$

(iii) $B \rightarrow \varepsilon \mid bB \mid aBb$

i. Ef fleiri a-tákn eru í strengnum en b-tákn, þá er valið „aA", en a-táknið tryggir að a-táknin verða fleiri.

Ef fleiri b-tákn eru í strengnum þá er valið „Bb", en b-táknið tryggir að b-táknin verða fleiri.

ii. „aAb" tryggir að ef bæta skal við b-tákni, þá helst alltaf meiri fjöldi a-tákna.

iii. Það sama gildir um „aBb", nema að þá er tryggt að b-táknin séu fleiri.

4.

Málinu má lýsa sem

$$L = \{ IF^n \; ELSE^m \mid n > 0 \text{ og } n \geq m \}.$$

Málið L hefur málfræðina

$G = (\{S, A\}, \{IF, ELSE\}, P, S)$,

þar sem reglur P:

$$S \longrightarrow IF \; A \mid IF \; A \; ELSE$$

$$A \longrightarrow S \mid \varepsilon \, .$$

Setningarnar í Java geta verið

eitt eða fleiri IF,
eitt eða fleiri IF-ELSE,
eða blanda af báðu.

5.

a.

$$G_A = (\{S, A, B\}, \{a, b\}, P, S)$$

med regler P:

$$S \to A \mid B$$

$$A \to aA \mid B$$

$$B \to bBc \mid \varepsilon$$

b.

$$G_B = (\{S, C, D\}, \{a, b\}, P, S)$$

med regler P:

$$S \to CD$$

$$C \to aCb \mid \varepsilon$$

$$D \to cD \mid \varepsilon$$

c.

Í sniðmáli $A \cap B$ þarf hver strengur að uppfylla kröfur A sem og kröfur B.

Skilyrðin sem strengirnir þarf því að uppfylla eru:

i. a- og b-táknin skulu vera í veldi $x \geq 0$.

ii. b- og c- táknin skulu vera í sama veldi.

iii. a- og b-táknin skulu vera í sama veldi.

Af skilyrðunum ii. og iii. má því sjá að a-, b-, og c-táknin þurfa öll að vera í sama veldi.

Því má lýsa máli $A \cap B$ sem:

$$A \cap B = \{a^n b^n c^n \mid n \geq 0\}.$$

Við vitum, samkvæmt fyrirlestri, að málið $\{a^n b^n c^n \mid n \geq 0\}$ er ekki samhengisfrjálst.

Þar sem A og B eru samhengisfrjáls, en sniðmengi $A \cap B$ er ekki samhengisfrjálst, þá höfum við sýnt fram á að samhengisfrjáls mál eru ekki lokuð með tilliti til sniðmáls.

6. Gerum ráð fyrir að A sé samhengisfrjálst mál og lát p vera gefið með dælusetningu.

Vel strenginn $s = a^p b^p b^p a^p$. Það er ljóst
① ② ③ ④
að $s \in A$ og $|s| \geq p$.

Athuga:

i. Til að skilyrðið $|vxy| \leq p$ sé uppfyllt, þarf skiptingin að vera þess eðlis að vxy nái aðeins yfir eitt af fjórum táknum s.
(①, ②, ③, ④)

ii. Ef vxy inniheldur annað af a-táknunum eða b-táknunum, eða bæði a- og b-tákn, þá inniheldur uv^2xy^2z runur af táknum sem eru ekki til staðar á hinum enda strengsins og því forsenda málsins, eða að $w = w^R$, brostin.

iii. Ef vxy inniheldur bæði b-táknin (②, ③), þá fylgir aukning á í óhjákvæmilega ójafn fjöldi a- og b-tákna og því er strengurinn ekki lengur í A.

Þetta dekkar öll tilfelli og við höfum þannig sýnt fram á mótsögn og A er því ekki reglulegt mál.