

Forstrengstré

Bergur Snorrason

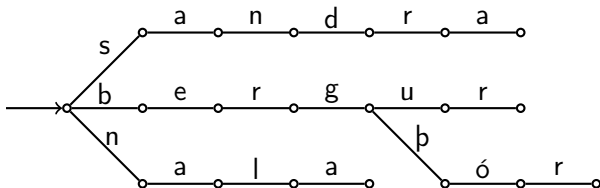
13. apríl 2023

- ▶ Við segjum að ekki tómt rótartré $T = (V, E)$ ásamt vörpun $\tau: E \rightarrow \Sigma$, þar sem Σ er eitthvert endanlegt mengi, sé *forstrengstré* (e. *prefix tree*) ef fyrir leggi e_1 og e_2 gildir $\tau(e_1) \neq \tau(e_2)$ ef leggirnir liggja úr sama hnút.
- ▶ Við köllum Σ *stafrófið* okkar og stökin í Σ köllum við *bókstafi*.
- ▶ Með öðrum orðum er τ átæk úr hverjum hnút eða, með enn öðrum orðum, engir leggir úr sama hnút geta verið merktir með sama bókstak.
- ▶ Forstrengstré eru oft kölluð *trie*.
- ▶ Við segjum að strengur s sé í trénu ef það er til hnútur v í trénu þannig að

$$s = \tau(e_1) \dots \tau(e_k)$$

þar sem e_1, \dots, e_k eru leggirnir á einfalda veginum frá rót til v , í réttri röð.

- ▶ Það er mjög algengt að geyma aukagögn í hnútunum í trénu.
- ▶ Skoðum dæmi um forstrengstré sem hefur engin aukagögn í hnútunum.



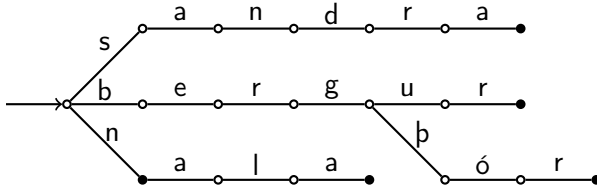
► Dæmi um strengi í trénu eru:

- „sandra”,
- „nala”,
- „bergur”,
- „bergþór”,
- „san” og
- „” (tómi strengurinn)

- ▶ Algengt er að merkja suma hnúta sem *lokahnúta*.
- ▶ Þetta eru dæmi um aukagögn sem við geymum í hnútum.
- ▶ Við segjum að strengur s sé í trénu ef það er til **lokahnútur** v í trénu þannig að

$$s = \tau(e_1) \dots \tau(e_k)$$

þar sem e_1, \dots, e_k eru leggirnir á einfalda veginum frá rót til v , í réttri röð.



- ▶ Strengirnir í trénu eru:
 - ▶ „sandra”,
 - ▶ „nala”,
 - ▶ „bergur”,
 - ▶ „bergþór” og
 - ▶ „n”
- ▶ Hvaða breytingu þurfum við að framkvæma ef við viljum að tómi strengurinn sé í trénu?

- ▶ Hvað gerum við ef við viljum bæta streng við forstrengstré?
- ▶ Ef við viljum setja strenginn $s = s_1 s_2 \dots s_n$ í tréð T þá setjum við strenginn $s' = s_2 s_3 \dots s_n$ í hluttré T sem við lendum í ef við flygjum leggnum merktum s_1 .
- ▶ Takið eftir að það hluttré má vera tóm (með öðrum orðum er ekki leggur merktur s_1).
- ▶ Í því tilfalli stækkar tréð.
- ▶ Í lokinn merkjum við hnútinn sem við endum í sem endahnút.

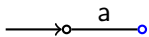


„api“

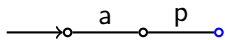
→o

„api“

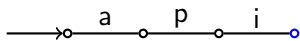




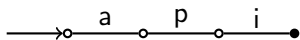
„pi”



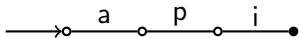
„I“



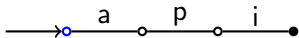
”
”



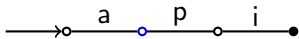
„apar”



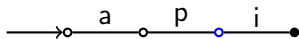
„apar”



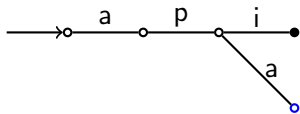
„par“



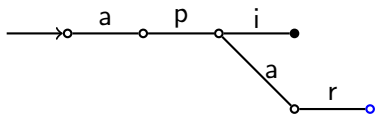
„ar“

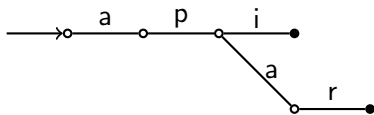


„r”

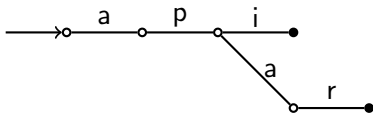


”
”

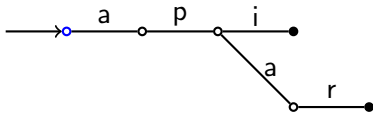




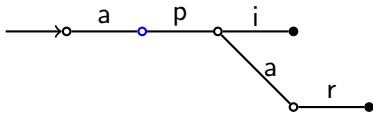
„apaköttur“



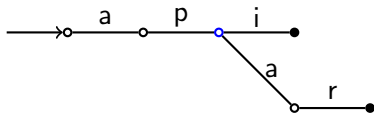
„apaköttur“



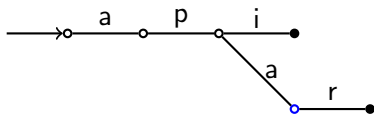
„paköttur“



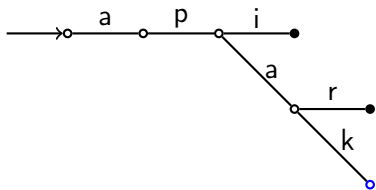
„aköttur“



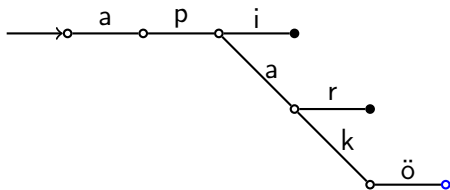
„köttur“



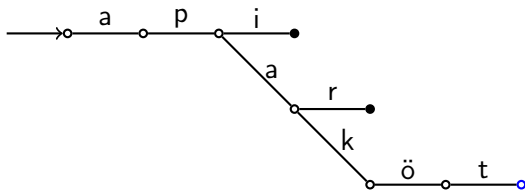
„öttur“



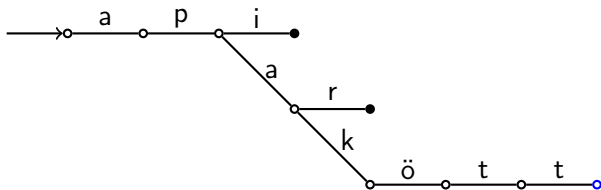
„ttur“



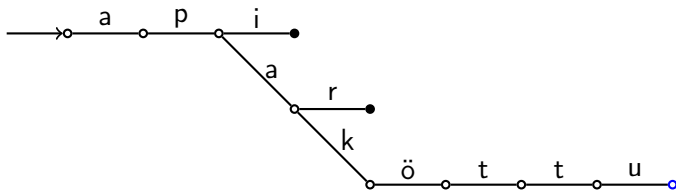
„tur“



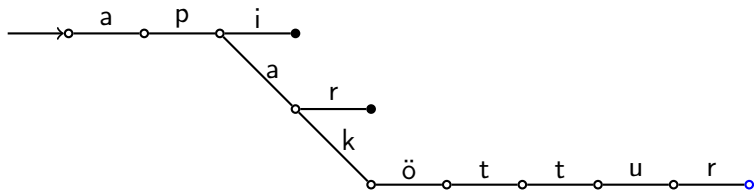
„ur“

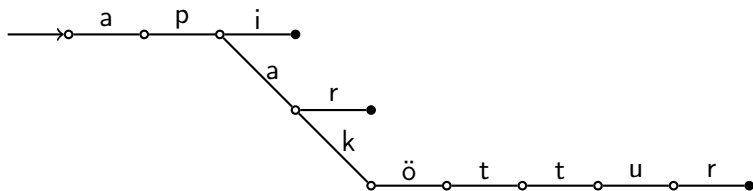


„r“

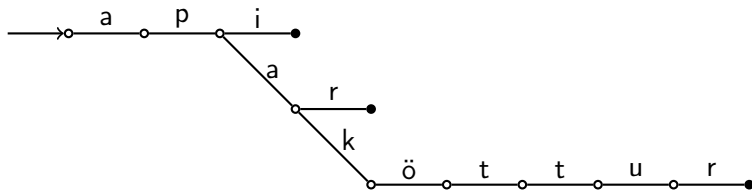


”
”

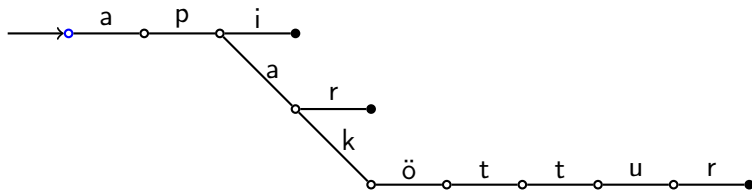




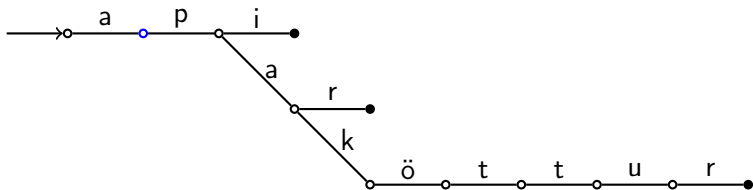
„altari“



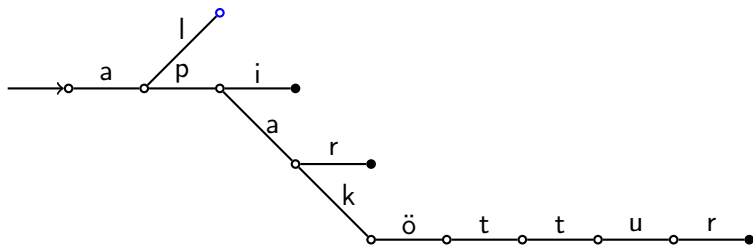
„altari“



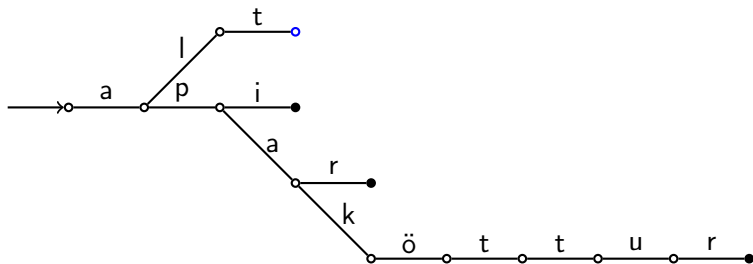
„Itari“



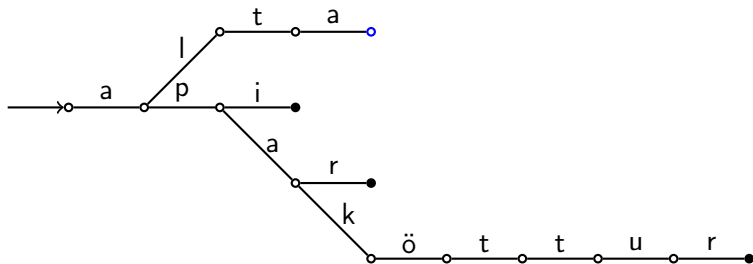
„tari“



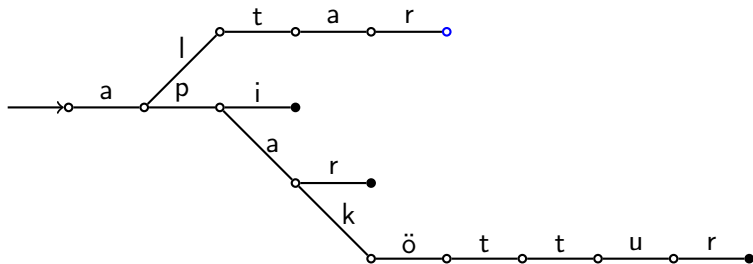
„ari“



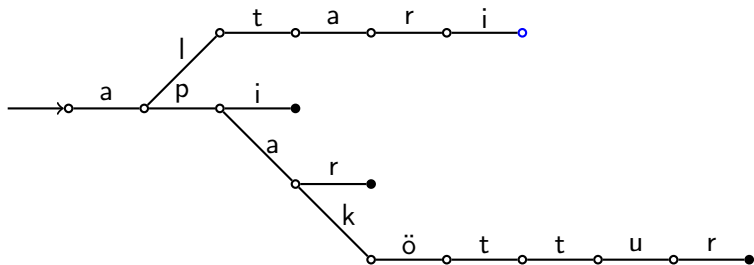
„ri“

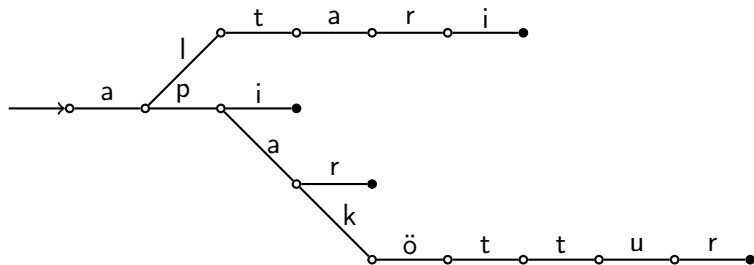


„I“

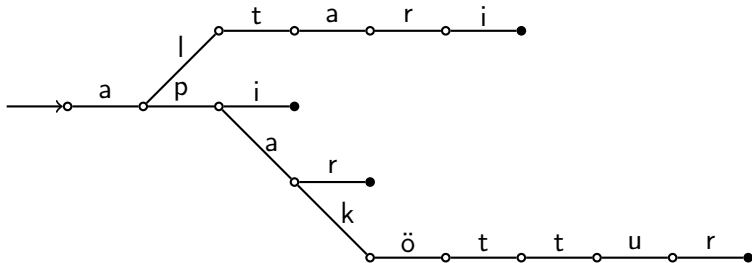


”
”

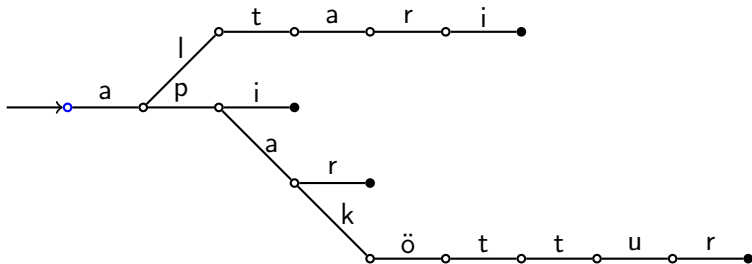




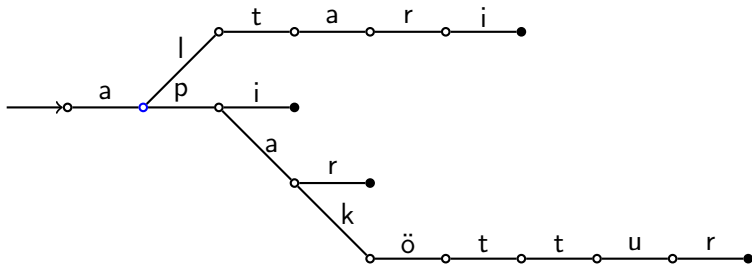
„apaspil”



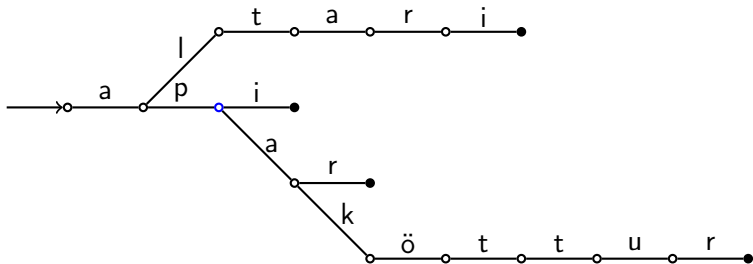
„apaspil”



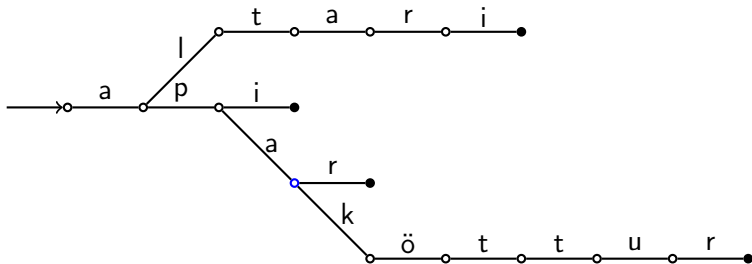
„paspil”



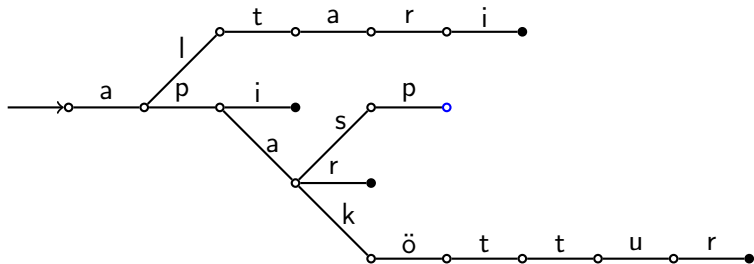
„aspil“



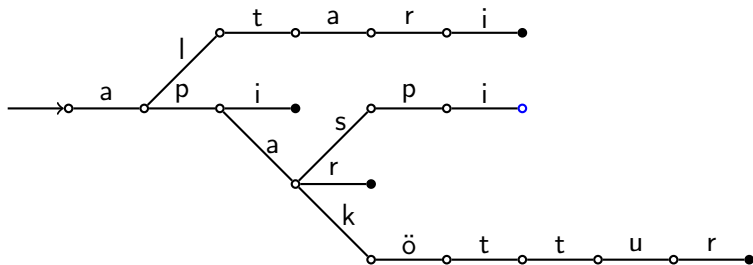
„spil“



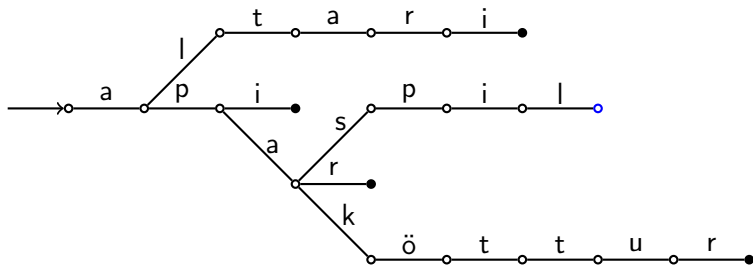
„il“

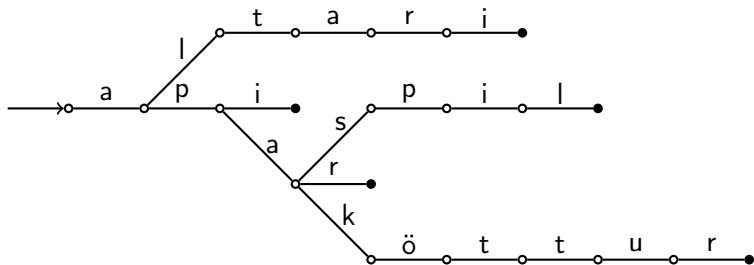


„l”

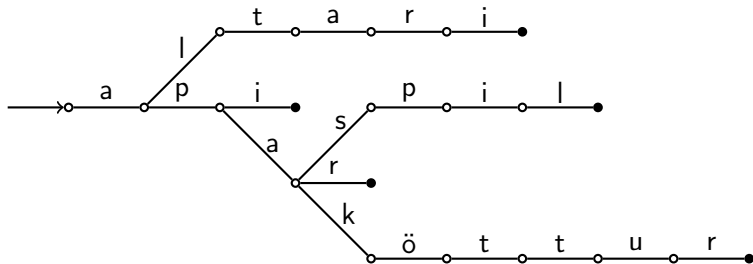


”
”

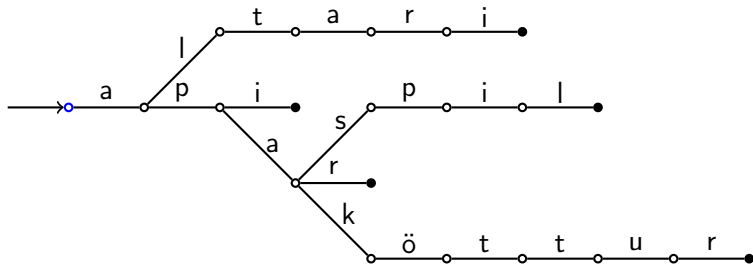




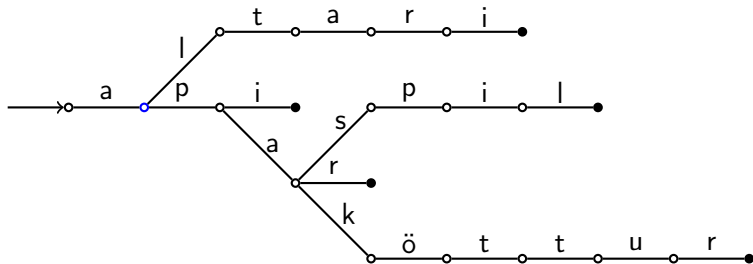
„altaristafla“



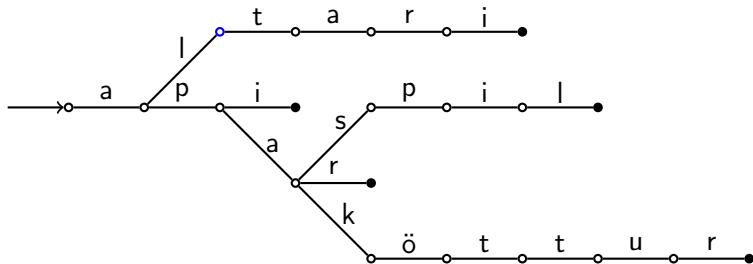
„altaristafla“



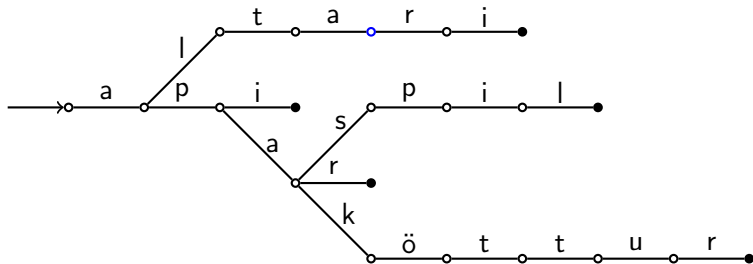
„ltaristafla”



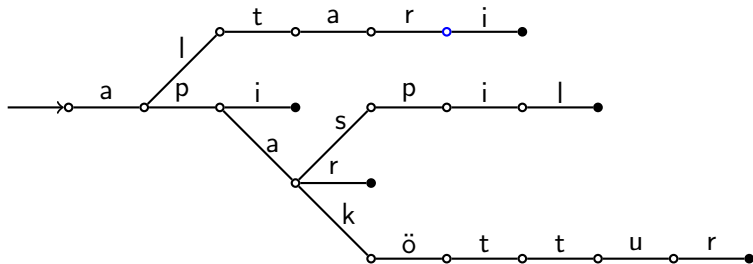
„taristafla“



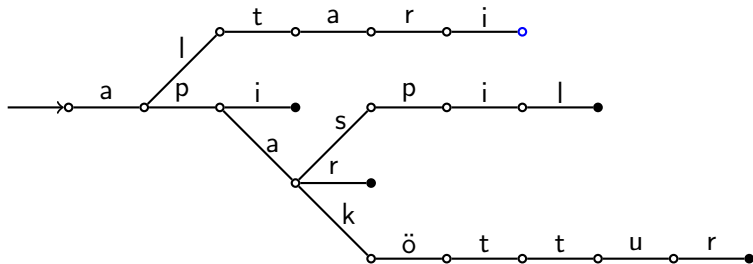
„ristafla“



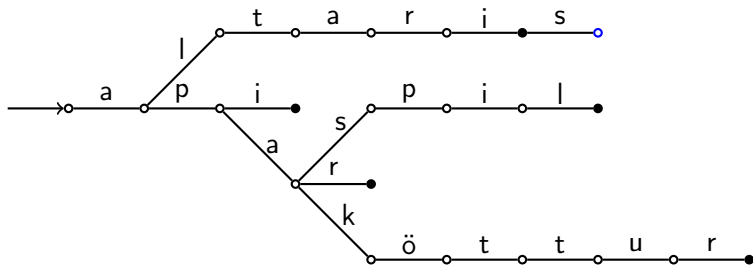
„istafla“



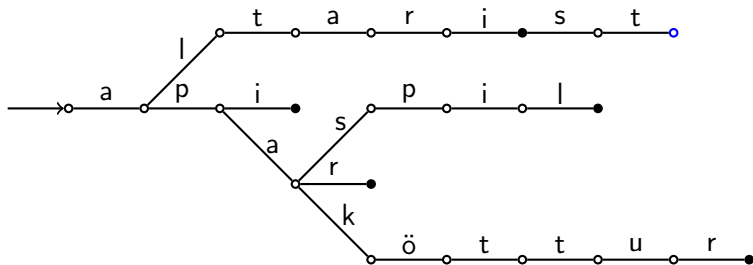
„stafla“



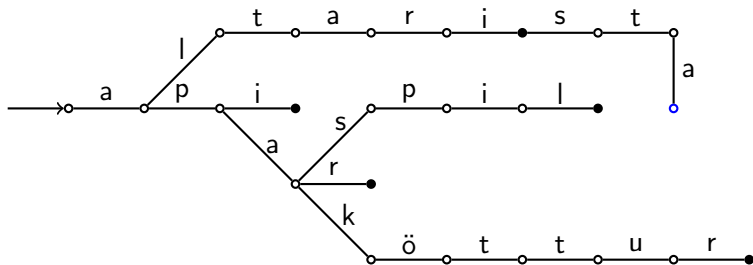
„tafla”



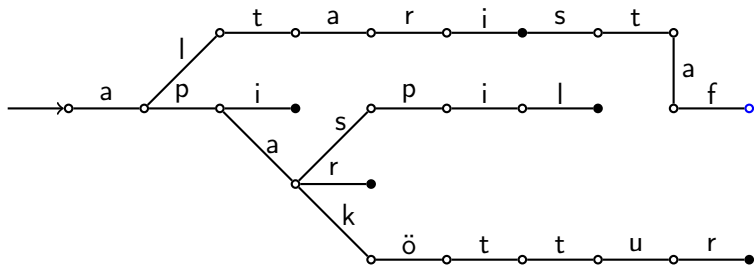
„afla”



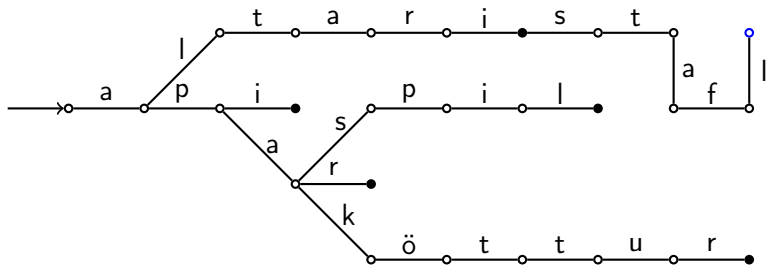
„fal“



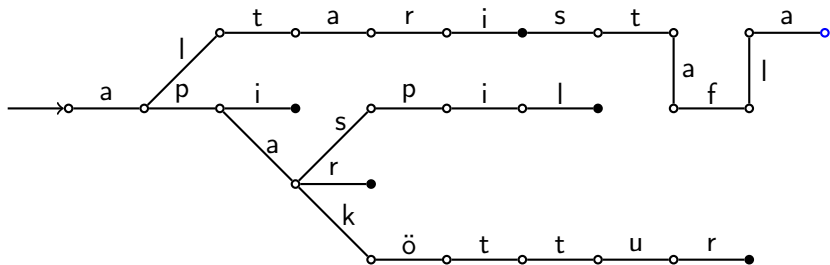
„la“

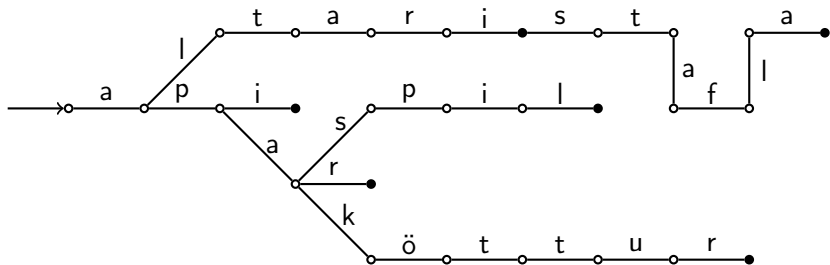


„a“

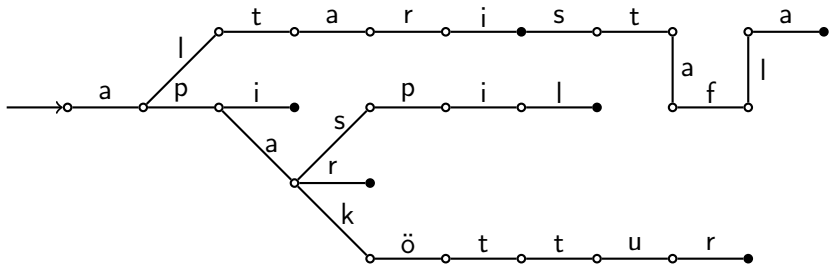


”
”

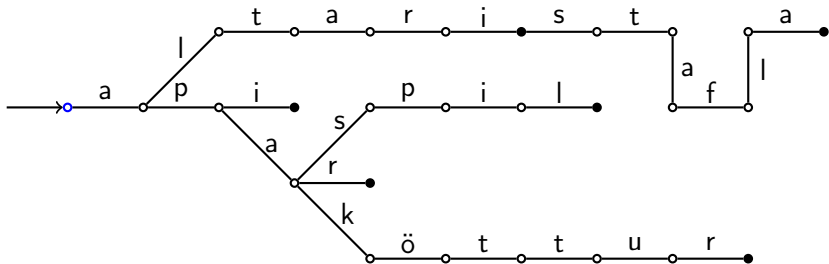




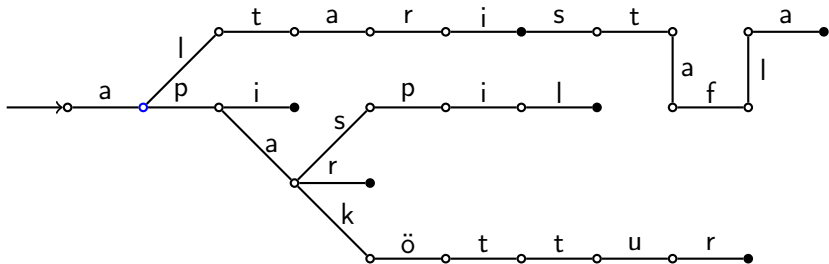
„altarisganga“



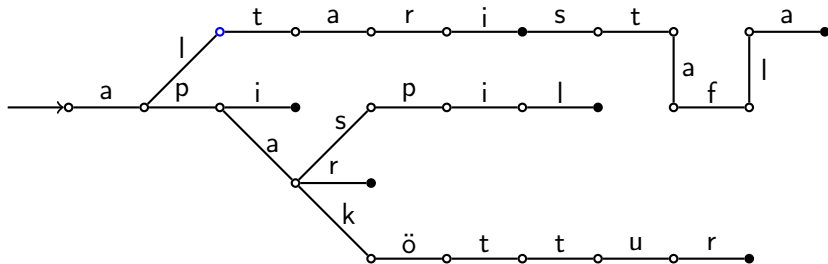
„altarisganga“



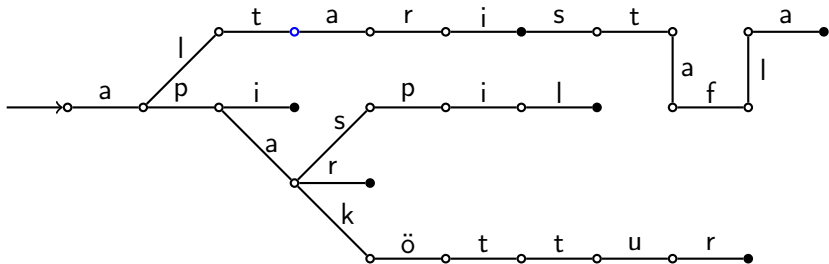
„Itarisganga“



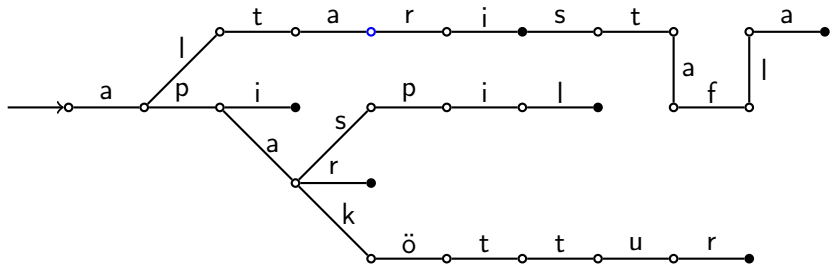
„tarisganga“



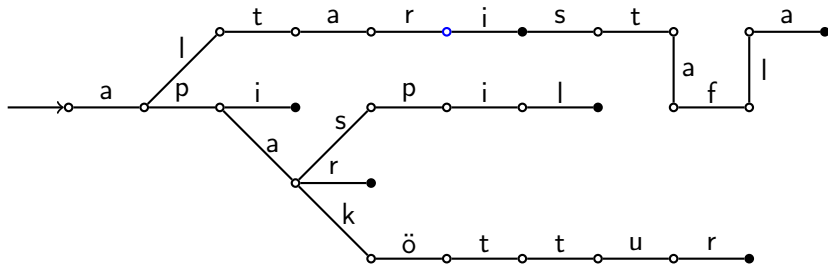
„arisganga“



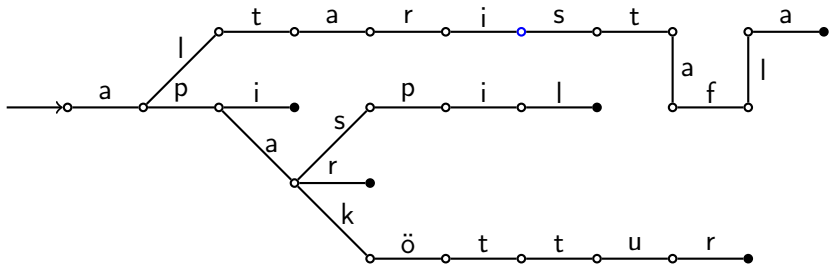
„risganga“



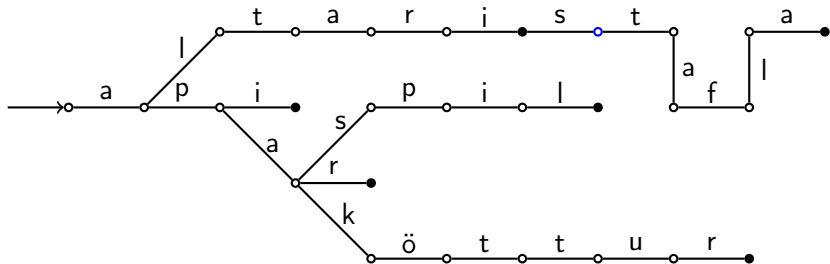
„isganga“



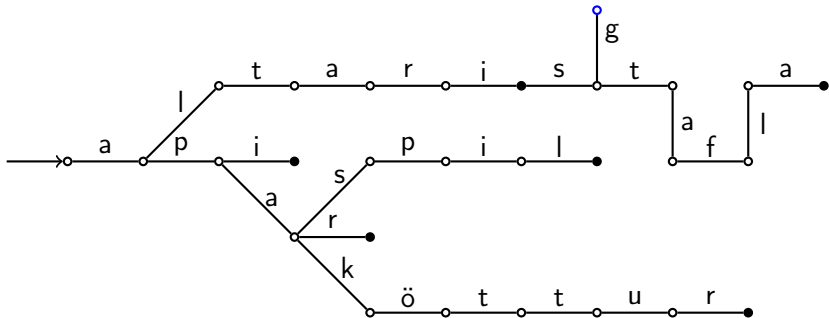
„sganga”



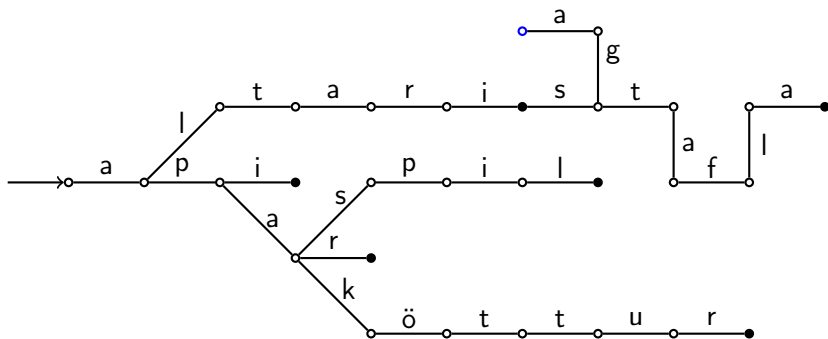
„ganga”



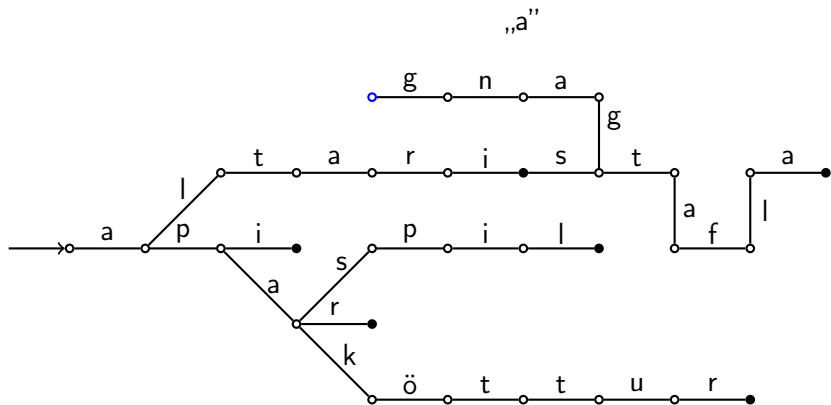
„anga”

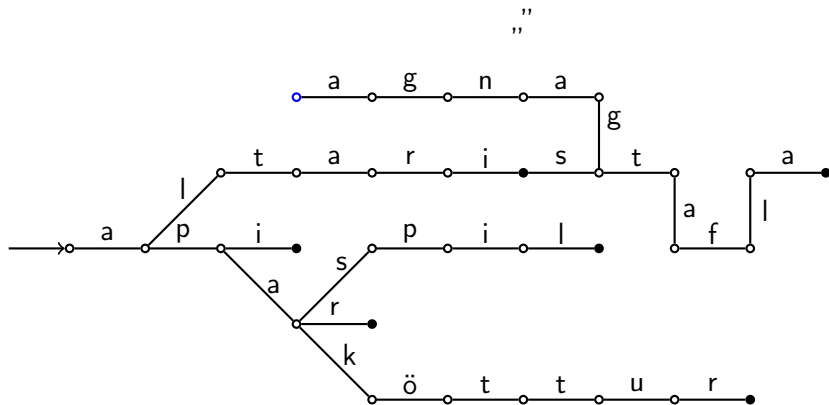


„nga”











- ▶ Til að útfæra forstrengstré munum við taka frá fylki af hnútum og úthluta þeim eftir þörf.
- ▶ Með þessa aðferð í huga er gott að vita hvað munum þurfa marga hnúta í heildina.
- ▶ Fyrir tómt forstrengstré þurfum við einn hnút.
- ▶ Ef við viljum bæta við streng þá þurfum við aldrei fleiri hnúta en lengd strengsins.
- ▶ Svo heildarfjöldi hnúta þarf að vera einum fleiri en samtals lengd allra strengjanna sem við setjum í forstrengstréð.

```

6 #define ALPHABET 26
7 #define MAXN 1000000
8 typedef struct { int t[ALPHABET], v; } trienode;
9 typedef struct { int s, r; trienode m[MAXN + 1]; } trie;
10 int trie_val(char c) { return c - 'a'; }
11 int trie_node(trie *t, int v)
12 {
13     for (int i = 0; i < ALPHABET; i++) t->m[t->s].t[i] = -1;
14     t->m[t->s].v = v;
15     return t->s++;
16 }
17 void trie_init(trie *t) { t->s = 0, t->r = trie_node(t, 0); }
18
19 void trie_insert(trie *t, char *s)
20 {
21     int h;
22     for (h = t->r; *s; h = t->m[h].t[trie_val(*s++)])
23         if (t->m[h].t[trie_val(*s)] == -1)
24             t->m[h].t[trie_val(*s)] = trie_node(t, 0);
25     t->m[h].v = 1;
26 }

```

- ▶ Ef við viljum setja strengi af samtals lengd n í forstrengstréð þá tekur það $\mathcal{O}(|\Sigma| \cdot n)$, þar sem $|\Sigma|$ er stærðin á stafrófinu.
- ▶ Takið eftir að ef stafrófið er takmarkað þá er tímaflækjan línuleg.

