Forstrengstré

Bergur Snorrason

12. apríl 2023

- Við segjum að ekki tómt rótartré $T = (V, E)$ ásamt vörpun $\tau \colon E \to \Sigma$, 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 segjum að ekki tómt rótartré $T = (V, E)$ ásamt vörpun $\tau\colon E \to \Sigma$, 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*.

- Við segjum að ekki tómt rótartré $T = (V, E)$ ásamt vörpun $\tau \colon E \to \Sigma$, 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.

- Við segjum að ekki tómt rótartré $T = (V, E)$ ásamt vörpun $\tau \colon E \to \Sigma$, 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 ekki tómt rótartré $T = (V, E)$ ásamt vörpun $\tau\colon E \to \Sigma$, 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

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

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

- Við segjum að ekki tómt rótartré $T = (V, E)$ ásamt vörpun $\tau \colon E \to \Sigma$, 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

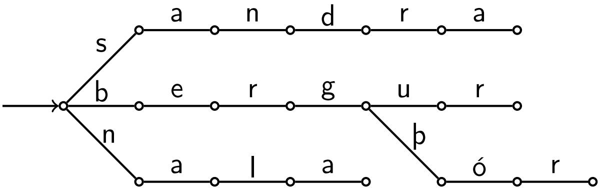
  $$s = \tau(e_1) \ldots \tau(e_k)$$

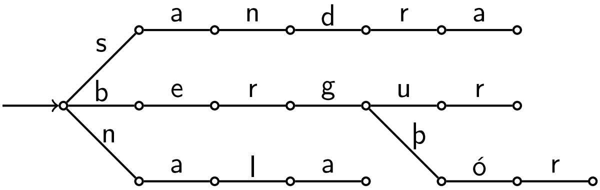
  þar sem $e_1, \ldots e_k$ eru leggirnir á einfalda veginum frá rót til v, í réttri röð.
- Það er mjög algengt að geyma aukagögn í hnútunum í trénu.

- Við segjum að ekki tómt rótartré $T = (V, E)$ ásamt vörpun $\tau \colon E \to \Sigma$, 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

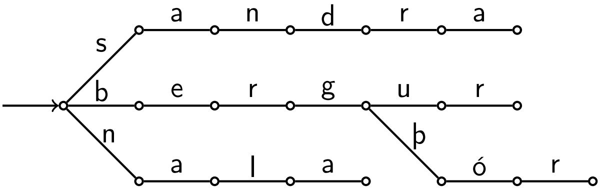
$$s = \tau(e_1) \ldots \tau(e_k)$$

  þar sem $e_1, \ldots e_k$ 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.

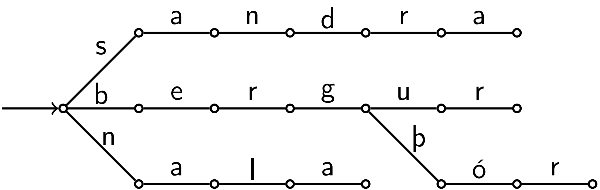

a n d r a
s
þ e r g u r
n þ
a l a ó r



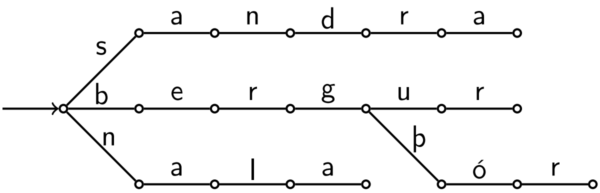
- Dæmi um strengi í trénu eru:



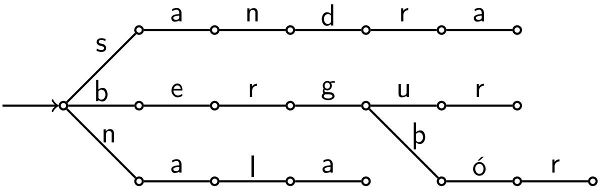
- Dæmi um strengi í trénu eru:
  - „sandra",



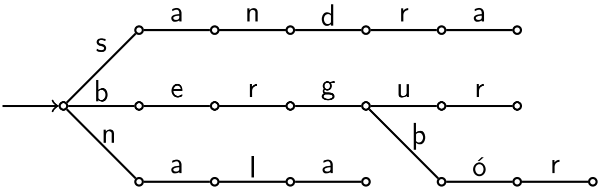
- Dæmi um strengi í trénu eru:
  - „sandra",
  - „nala",



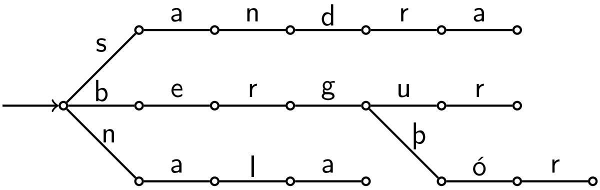
- Dæmi um strengi í trénu eru:
  - „sandra“,
  - „nala“,
  - „bergur“,



- Dæmi um strengi í trénu eru:
  - „sandra",
  - „nala",
  - „bergur",
  - „bergþór",



- Dæmi um strengi í trénu eru:
  - „sandra",
  - „nala",
  - „bergur",
  - „bergþór",
  - „san" og

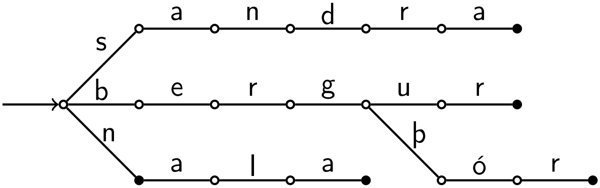


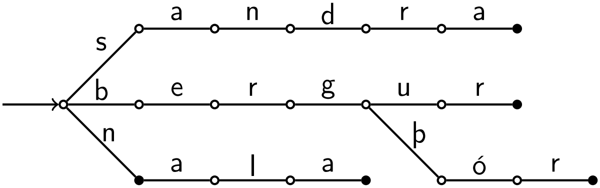
- Dæmi um strengi í trénu eru:
  - „sandra",
  - „nala",
  - „bergur",
  - „bergþór",
  - „san" og
  - „" (tómi strengurinn)

- Algengt er að merkja suma hnúta sem *lokahnúta*.

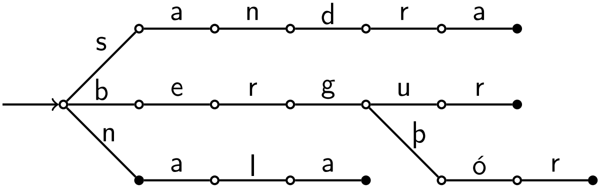
- Algengt er að merkja suma hnúta sem *lokahnúta*.
- Þetta eru dæmi um aukagögn sem við geymum í hnútum.

- 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 **loka**hnútur v í trénu þannig að
$$s = \tau(e_1)\ldots\tau(e_k)$$
þar sem $e_1, \ldots e_k$ eru leggirnir á einfalda veginum frá rót til v, í réttri röð.

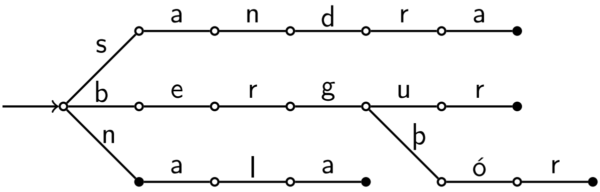

s a n d r a
b e r g u r
n a l a
þ ó r



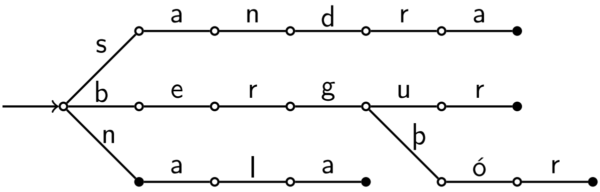
- Strengirnir í trénu eru:



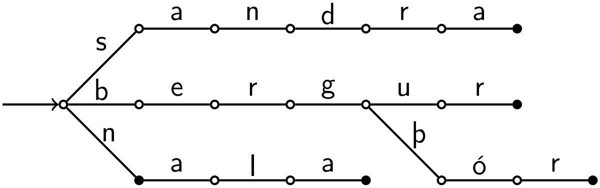
- Strengirnir í trénu eru:
  - „sandra",



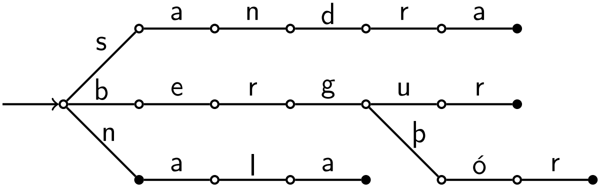
- Strengirnir í trénu eru:
  - „sandra“,
  - „nala“,



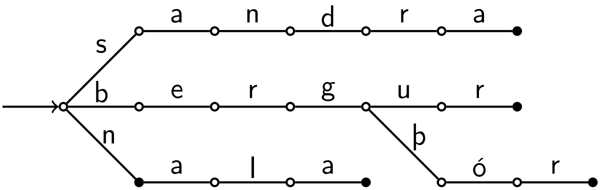
- Strengirnir í trénu eru:
  - „sandra",
  - „nala",
  - „bergur",



- Strengirnir í trénu eru:
  - „sandra",
  - „nala",
  - „bergur",
  - „bergþór" og



- Strengirnir í trénu eru:
  - „sandra",
  - „nala",
  - „bergur",
  - „bergþór" og
  - „n"



- 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?

- Hvað gerum við ef við viljum bæta streng við forstrengstré?
- Ef við viljum setja strenginn $s = s_1 s_2 \ldots s_n$ í tréð T þá setjum við strenginn $s' = s_2 s_3 \ldots s_n$ í hluttré T sem við lendum í ef við flygjum leggnum merktum s_1.

- 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 tómt (með öðrum orðum er ekki leggur merktur s_1).

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

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

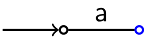
$\longrightarrow\circ$

„api”

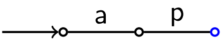
$\longrightarrow\!\circ$

„api"

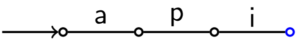
„pi”

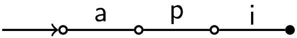


„i”

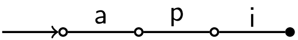


’’
,,

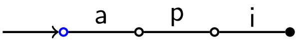

a
p
i


a
p
i

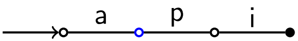
„apar''



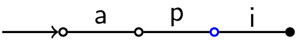
„apar”



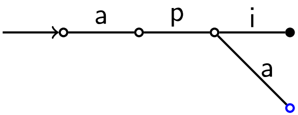
„par”



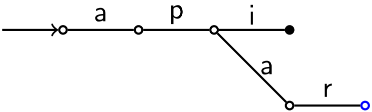
„ar”



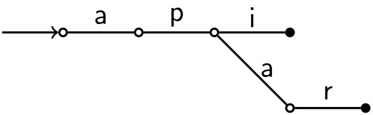
„r''



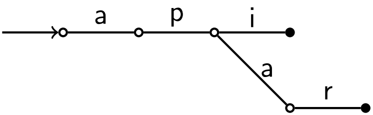
,,
,,



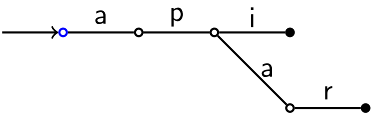
a
p
i
a
r


a
p
i
a
r

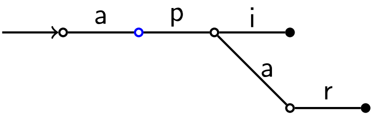
„apaköttur”



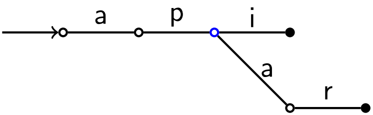
„apaköttur"



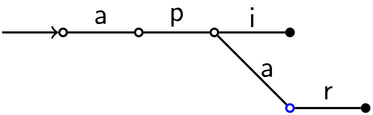
„pakötttur”



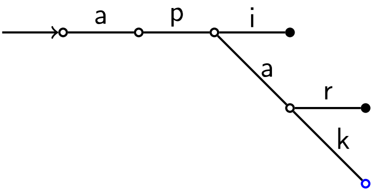
„aköttur”



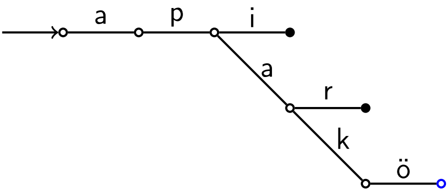
„köttur"



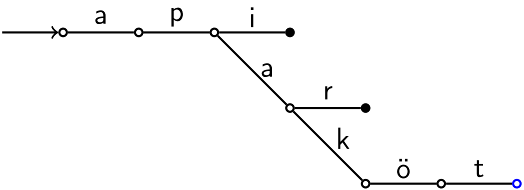
„öttur"



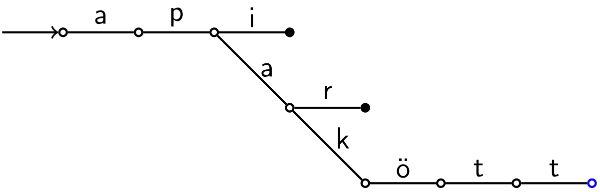
„ttur”



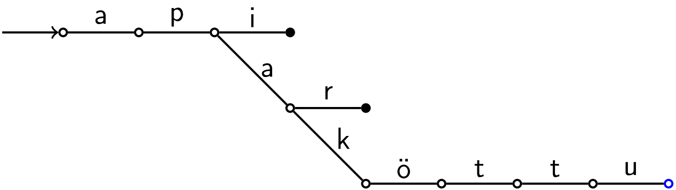
„tur”



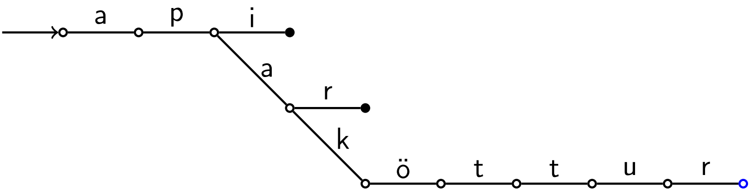
„ur"

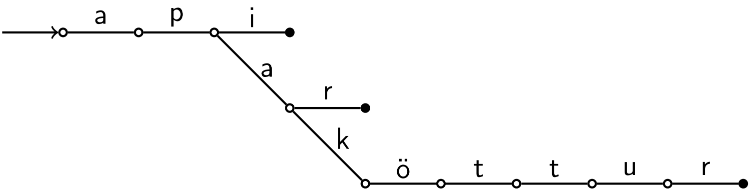


„r''

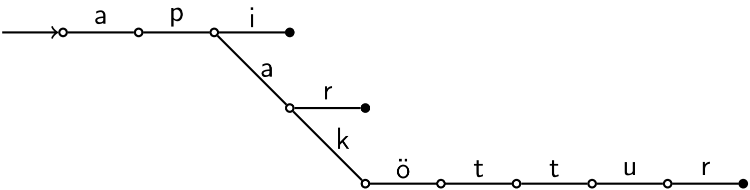


,''
,,

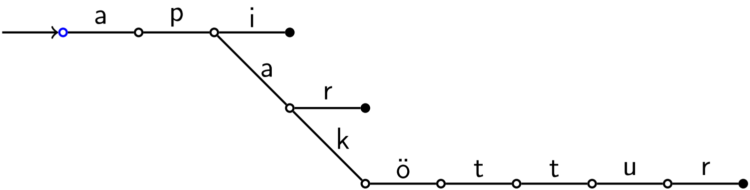

a
p
i
a
r
k
ö
t
t
u
r


a
p
i
a
r
k
ö
t
t
u
r

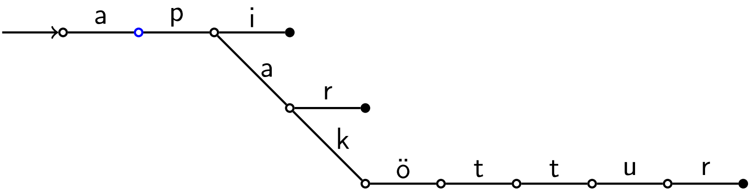
„altari"



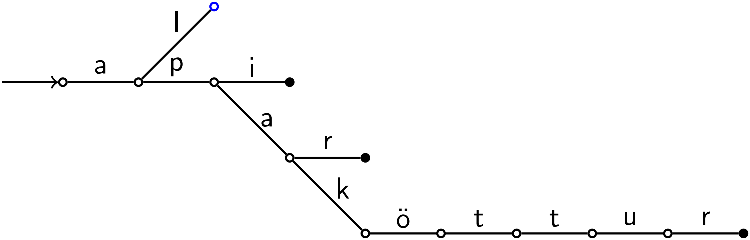
„altari"



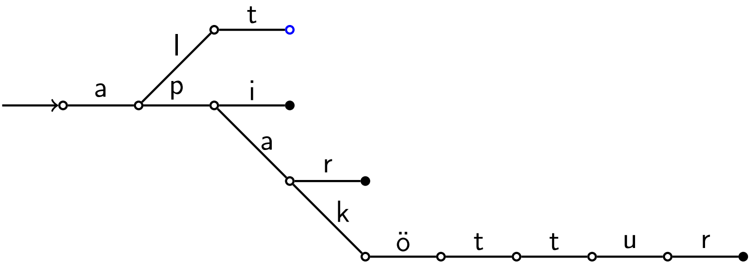
„Itari”



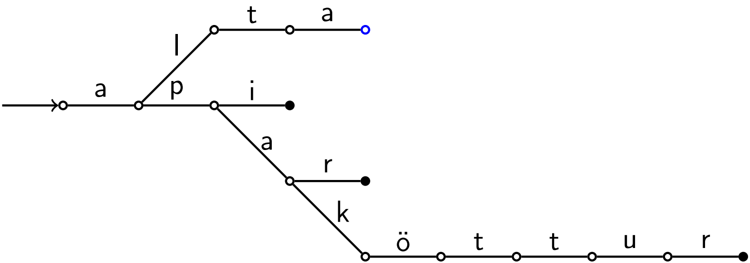
„tari"



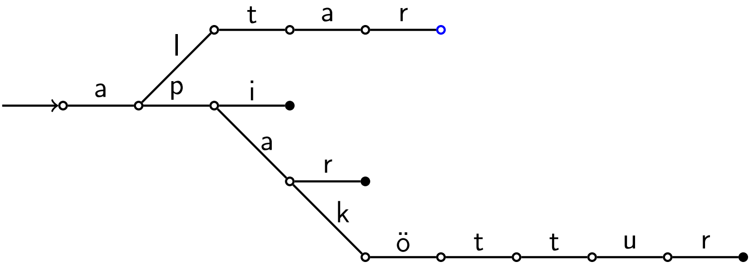
„ari”



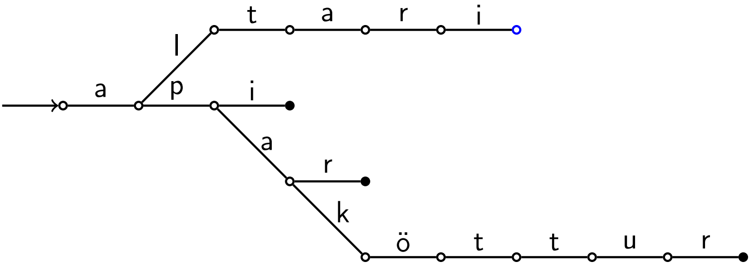
„ri''



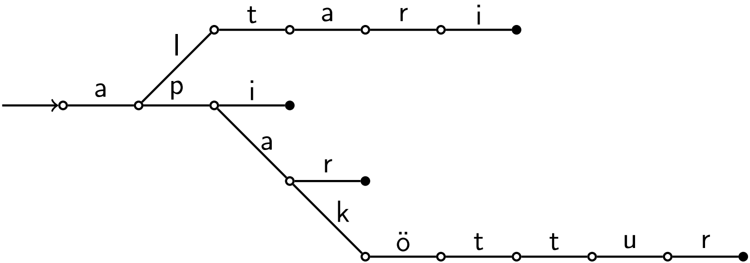
„i"



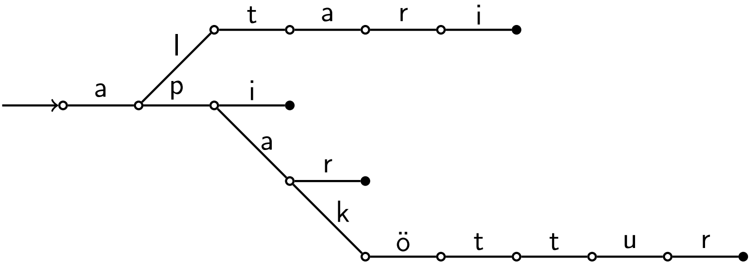
''
,,



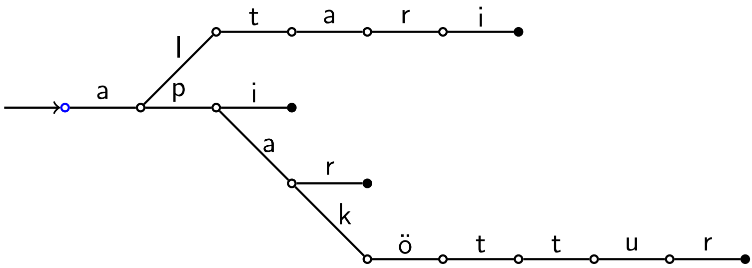
t a r i
l
a p i
a
r
k
ö t t u r


t
a
r
i
l
a
p
i
a
r
k
ö
t
t
u
r

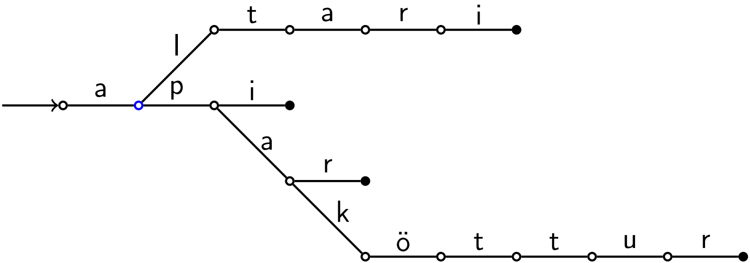
„apaspil"



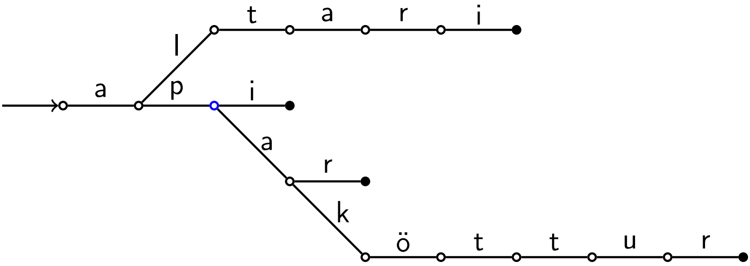
„apaspil”



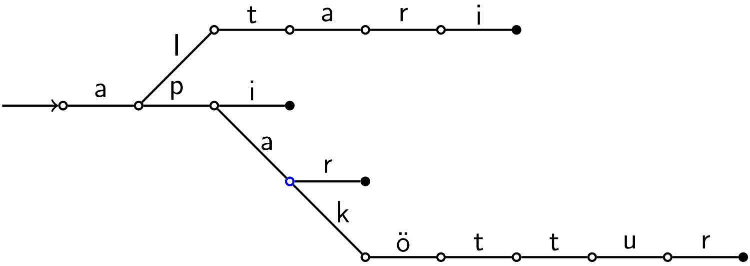
„paspil"



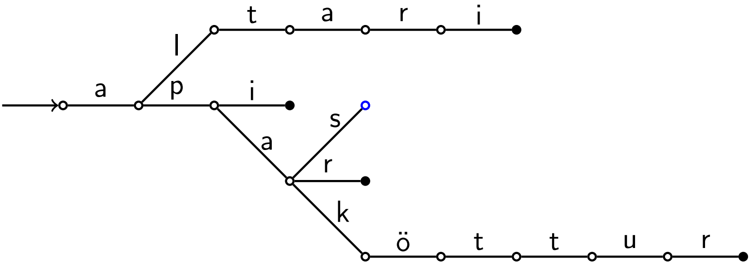
„aspil"



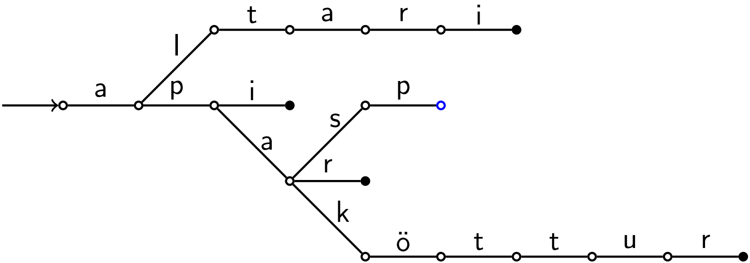
„spil''



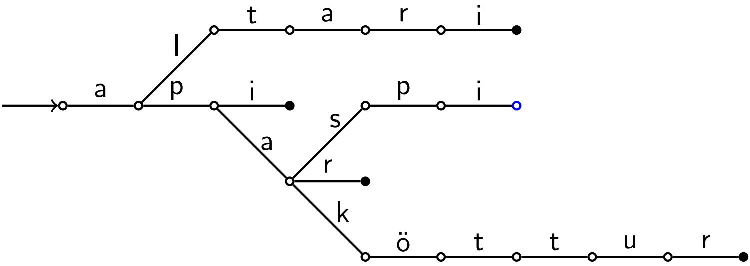
„pil”



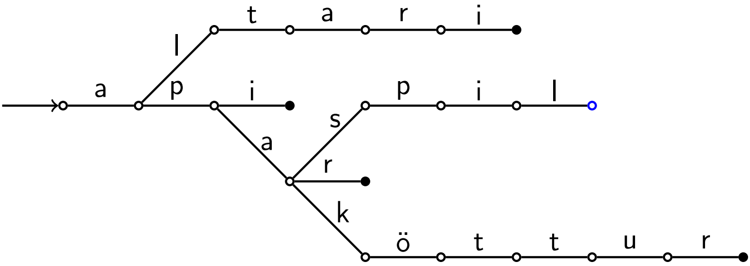
„il”

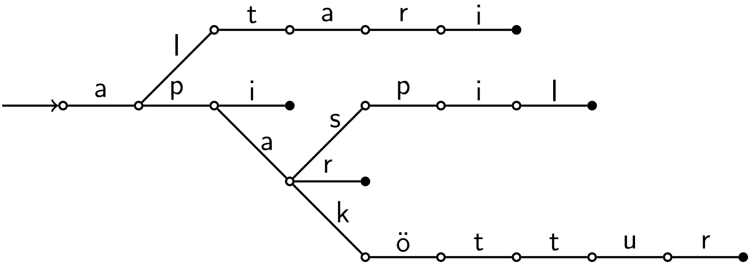


„l"

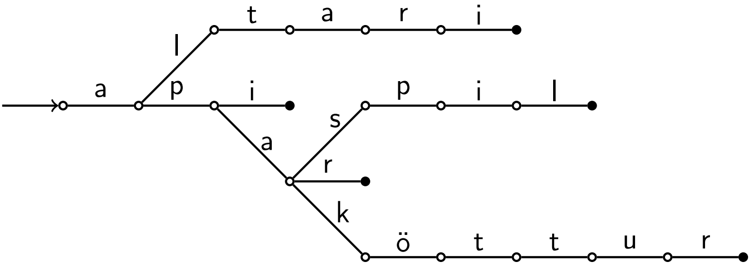


„”

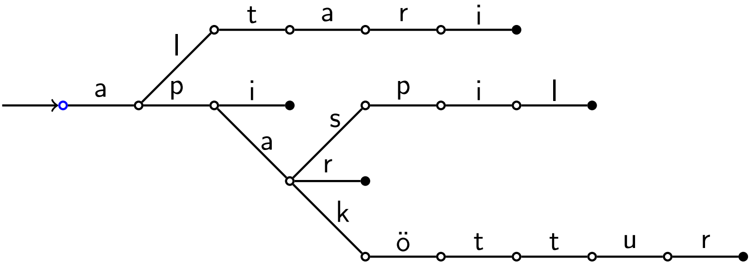

a
l
t a r i
p
i
a
s
p i l
r
k
ö t t u r


a
l
t
a
r
i
p
i
p
i
l
s
a
r
k
ö
t
t
u
r

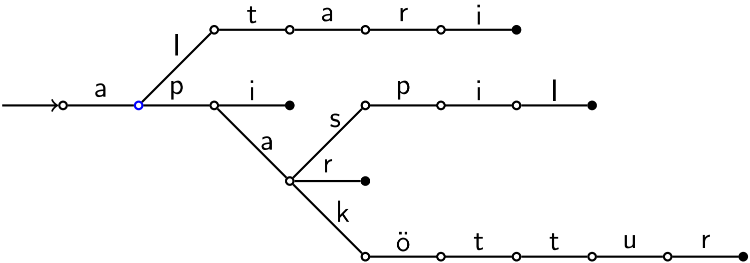
„altaristafla”



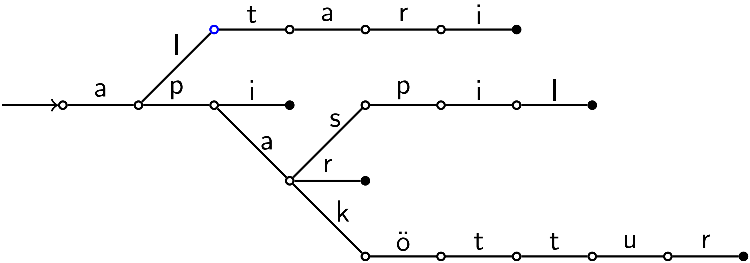
„altaristafla”



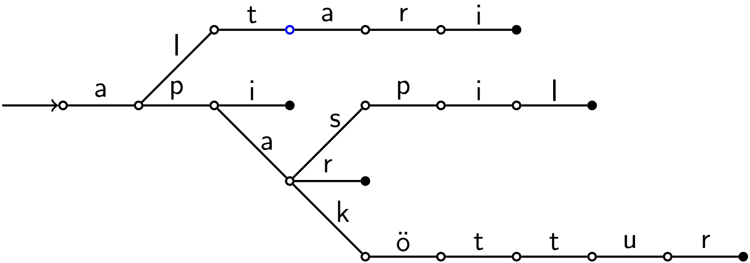
„ltaristafla”



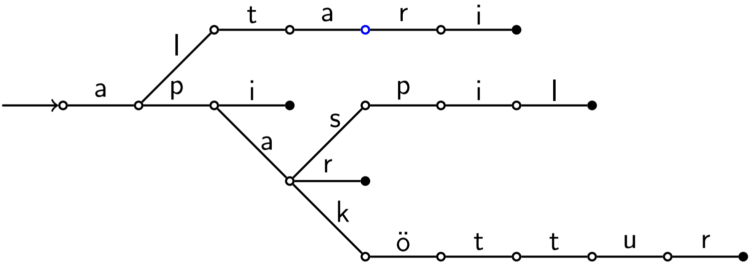
„taristafla"



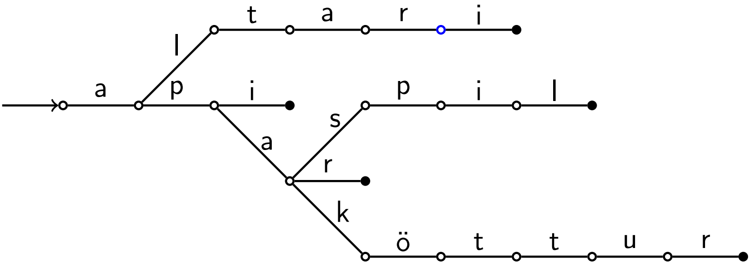
„aristafla"



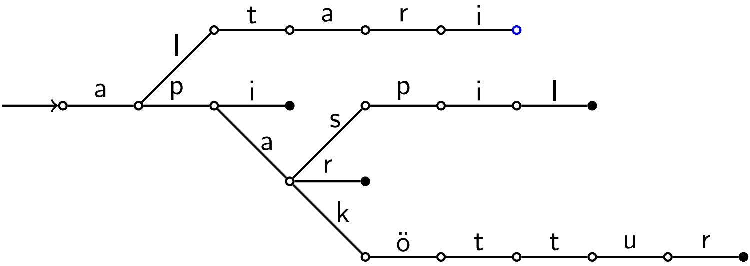
„ristafla”



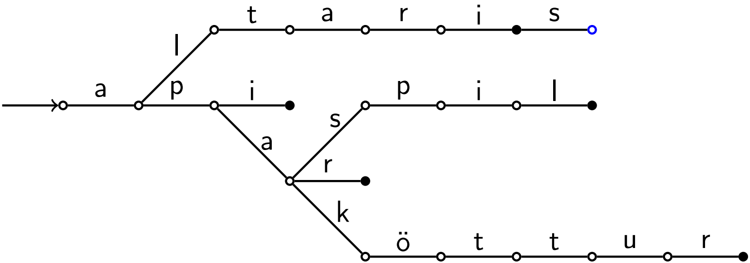
„istafla”



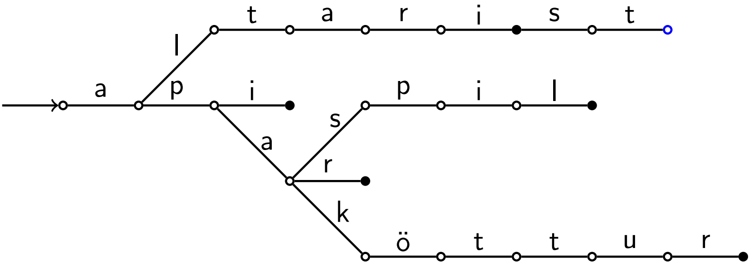
„stafla"



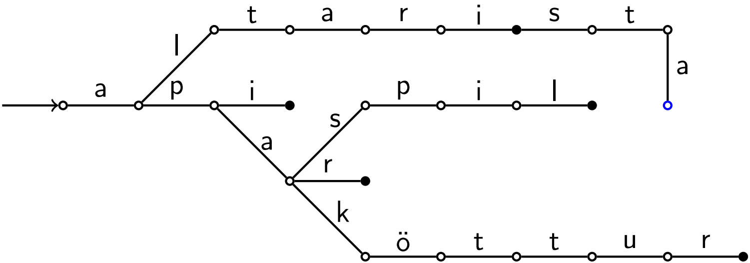
„tafla"



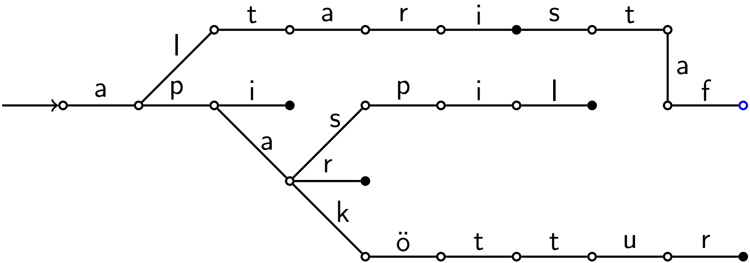
„afla”



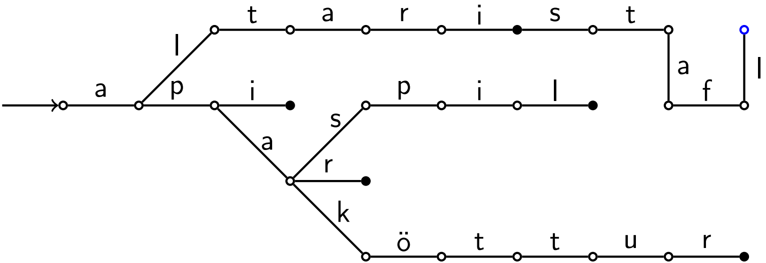
„fal”



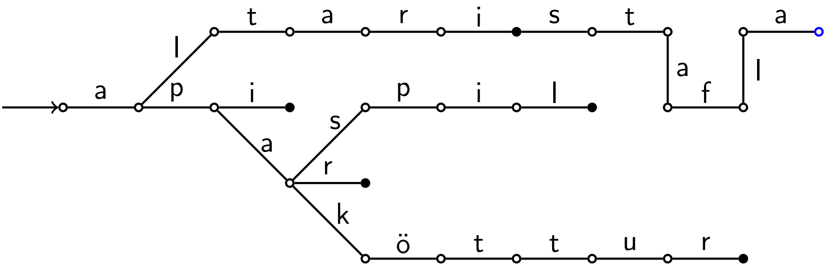
„la”

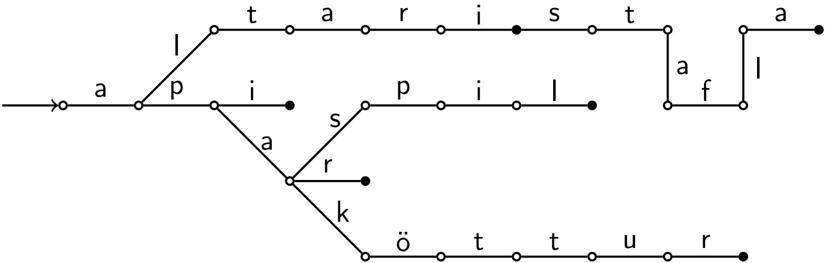


„a”

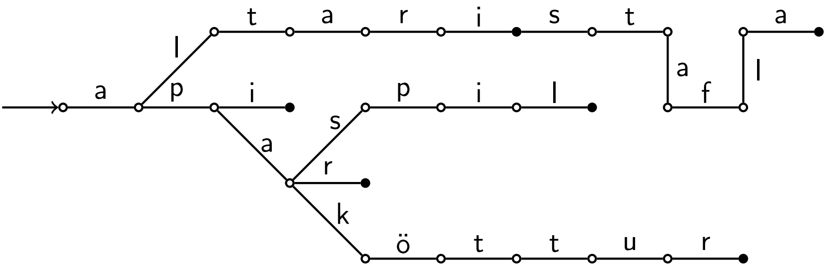


„"

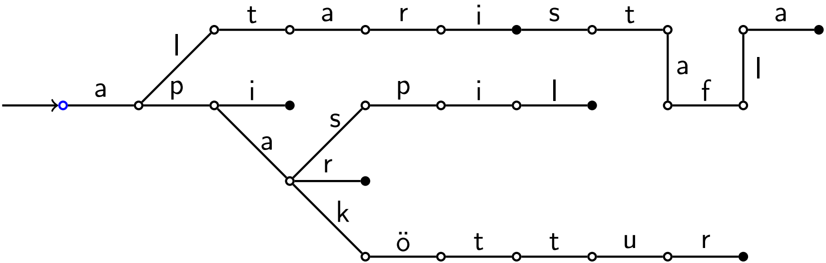

t a r i s t a
l
a f l
a p i p i l
s
a
r
k
ö t t u r


t a r i s t a
l a f l
a p i p i l
s a
r
k
ö t t u r

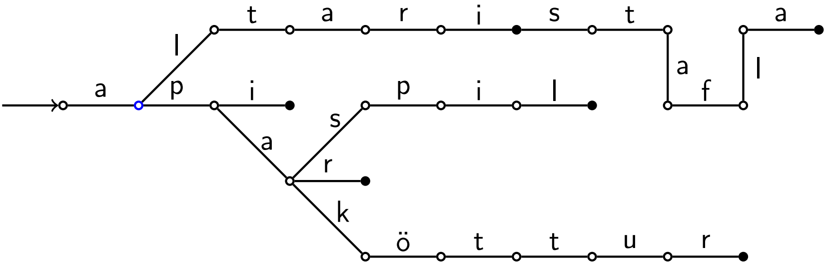
„altarisganga"



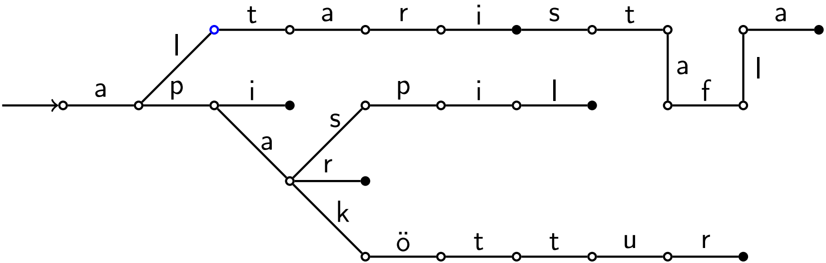
„altarisganga”



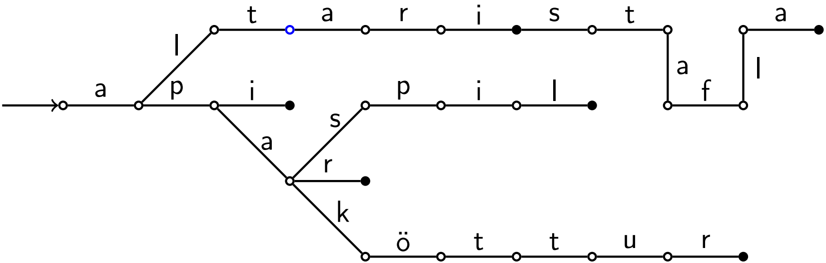
„ltarisganga"



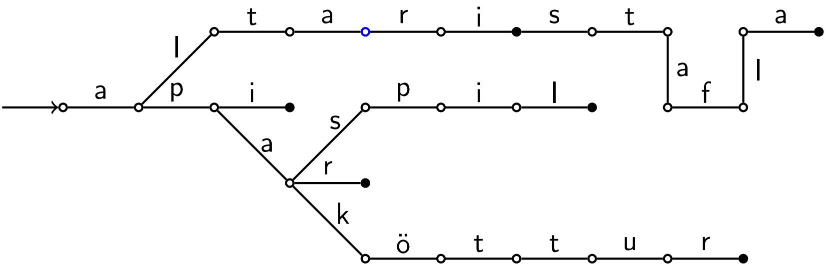
„tarisganga"



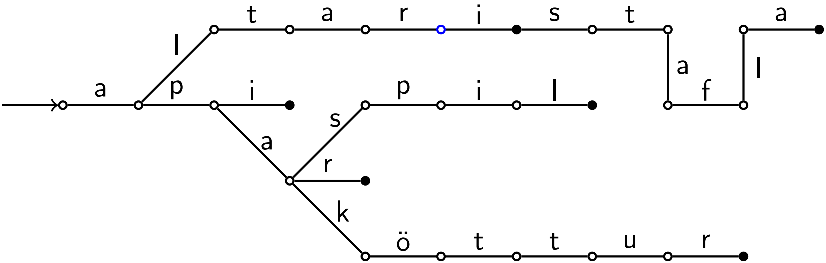
„arisganga"



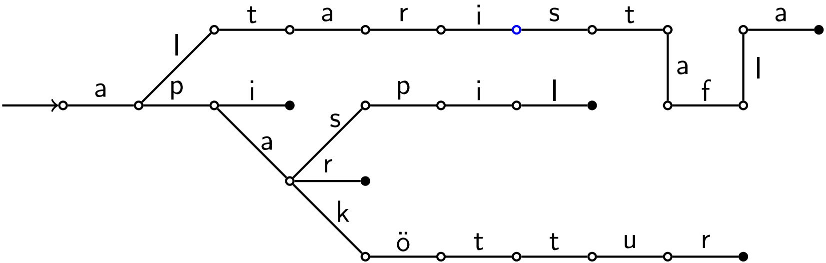
„risganga”



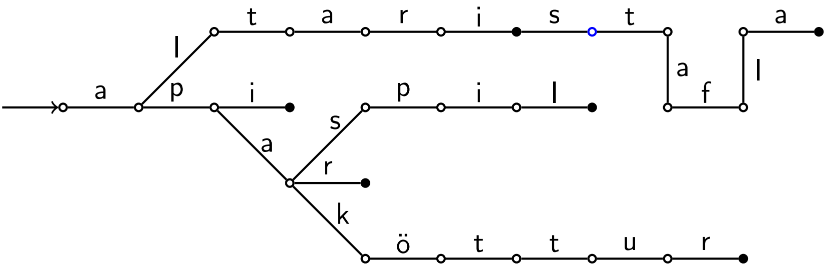
„isganga"



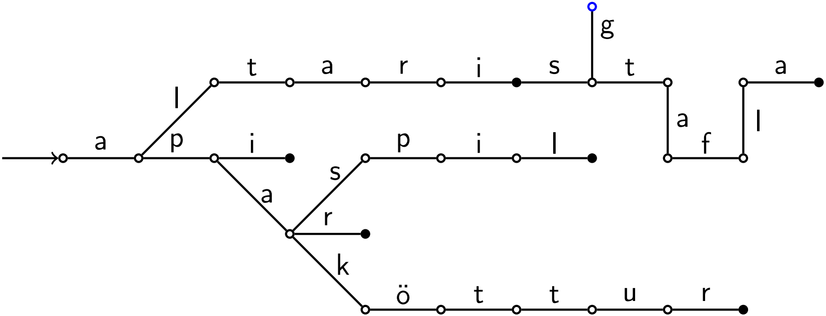
„sganga”



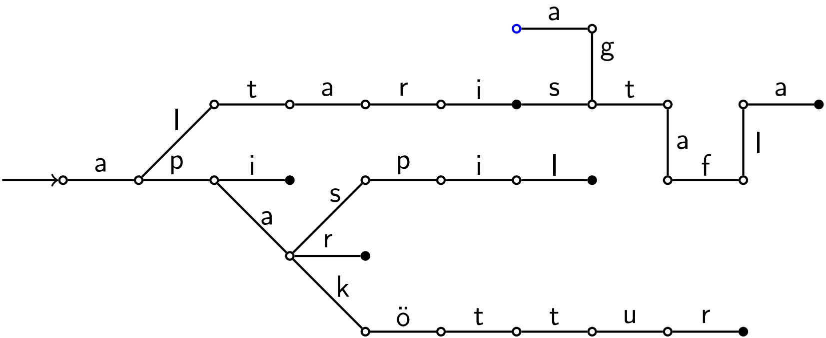
„ganga"



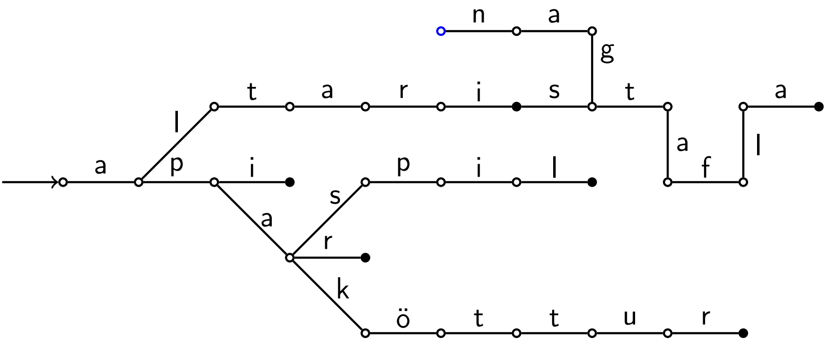
„anga"



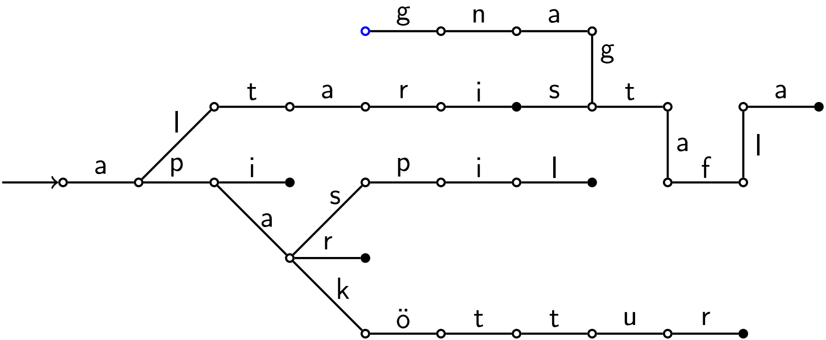
„nga"



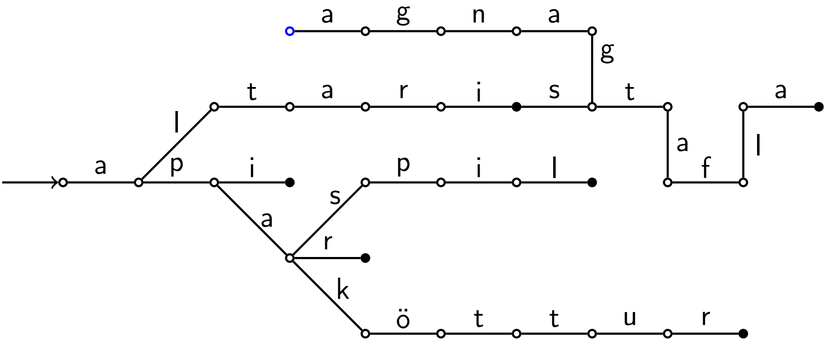
„ga”



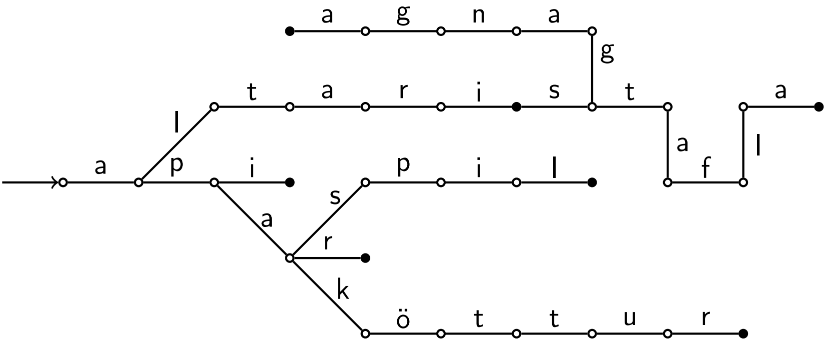
„a”



„"



a g n a
g
t a r i s t a
l a
l
a p p i l
i
a f
s
a
r
k
ö t t u r


a g n a
g
l t a r i s t a
a l
a p i p i l f
s
a
r
k
ö t t u r

- Til að útfæra forstrengstré munum við taka frá fylki af hnútum og úthluta þeim eftir þörf.

- 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.

- 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.

- 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.

- 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.

```c
#define ALPHABET 26
#define MAXN 1000000
typedef struct { int t[ALPHABET], v; } trienode;
typedef struct { int s, r; trienode m[MAXN + 1]; } trie;
int trie_val(char c) { return c - 'a'; }
int trie_node(trie *t, int v)
{
    for (int i = 0; i < ALPHABET; i++) t->m[t->s].t[i] = -1;
    t->m[t->s].v = v;
    return t->s++;
}
void trie_init(trie *t) { t->s = 0, t->r = trie_node(t, 0); }

void trie_insert(trie *t, char *s)
{
    int h;
    for (h = t->r; *s; h = t->m[h].t[trie_val(*s++)])
        if (t->m[h].t[trie_val(*s)] == -1)
            t->m[h].t[trie_val(*s)] = trie_node(t, 0);
    t->m[h].v = 1;
}
```

- Ef við viljum setja strengi af samtals lengd n í forstrengstréð þá tekur það $\mathcal{O}(\qquad)$, þar sem $|\Sigma|$ er stærðin á stafrófinu.

- 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.

- 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.