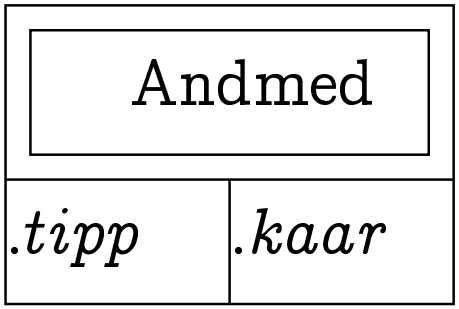
Suunatud graaf G koosneb

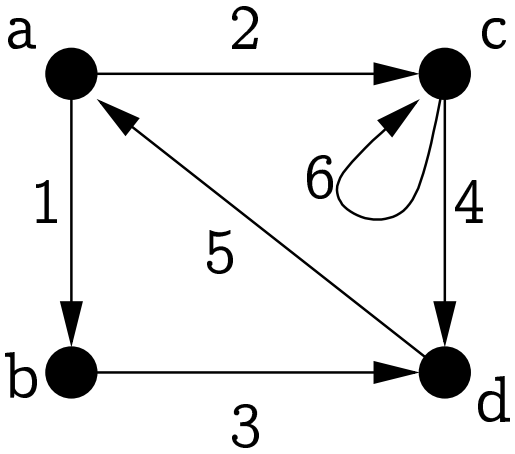
- *tippude* hulgast V;
- *servade* ehk *kaarte* hulgast E;
- *intsidentsusfunktsioonist* $\mathcal{E} : E \longrightarrow V \times V$, mis seab igale kaarele vastavusse tema alg- ja lõpptipu.

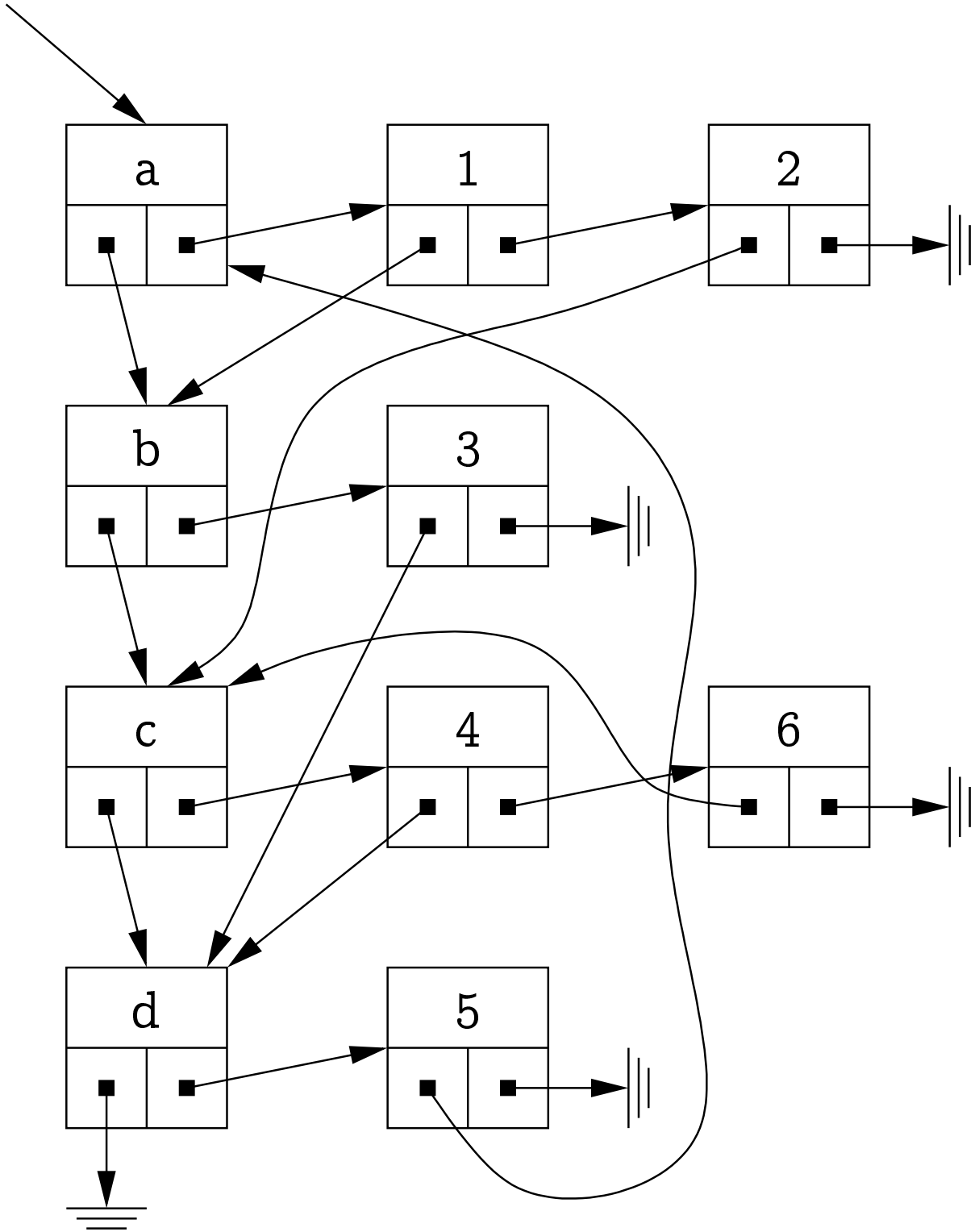
Suunamata serva kujutame üht ja teist pidi suunatud servade paarina.

Graafi võib mälus kujutada struktuurina, kus igale tipule ja igale servale vastab kirje



- Graafi tipud on lihtahelas välja *.tipp* kaudu; viit esimesele tipule on viit kogu graafile.
- Kõik mingist tipust algavad servad on lihtahelas välja *.kaar* kaudu; esimesele lülile selles ahelas viitab selle tipu väli *.kaar*.
- Servadele vastavates kirjetes viitab väli *.tipp* selle serva lõpptipule (vastavale kirjele).


a
2
c
1
6
4
5
b
3
d


a
1
2
b
3
c
4
6
d
5

Dünaamiline järjend: andmetüüp, mille väärtusvaruks on kirjete suvalise pikkusega järjendid.

Magasin: dünaamiline järjend, kus elemente lisatakse alati järjendi lõppu ja elemente võetakse järjendi lõpust.

Järjekord: dünaamiline järjend, kus elemente lisatakse alati järjendi lõppu ja elemente võetakse järjendi algusest.

Elemendi r lisamist järjendisse Q tähistame $Q \Leftarrow r$. Elemendi võtmist järjendist Q ja omistamist muutujale r tähistame $r \Leftarrow Q$.

Operatsioonid: Lisa 1, Lisa 2, Lisa 3, Võta, Lisa 4, Võta, Võta

Magasin:

Võetud:

Järjekord:

Võetud:

Operatsioonid: Lisa 2, Lisa 3, Võta, Lisa 4, Võta, Võta

Magasin: 1

Võetud:

Järjekord: 1

Võetud:

Operatsioonid: Lisa 3, Võta, Lisa 4, Võta, Võta

Magasin:

1	2

Võetud:

Järjekord:

1	2

Võetud:

Operatsioonid: Võta, Lisa 4, Võta, Võta

Magasin:

1	2	3

Võetud:

Järjekord:

1	2	3

Võetud:

Operatsioonid: Lisa 4, Võta, Võta

Magasin:

1	2

Võetud: 3

Järjekord:

2	3

Võetud: 1

Operatsioonid: Võta, Võta

Magasin:

1	2	4

Võetud: 3

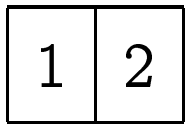
Järjekord:

2	3	4

Võetud: 1

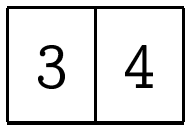
Operatsioonid: Võta

Magasin:



Võetud: 3, 4

Järjekord:



Võetud: 1, 2

Operatsioonid:

Magasin: 1

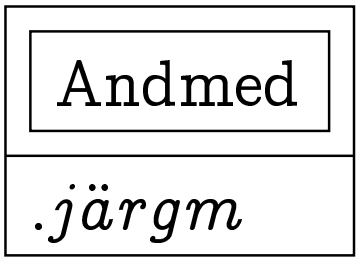
Võetud: 3, 4, 2

Järjekord: 4

Võetud: 1, 2, 3

Magasini võib realiseerida dünaamilise andmestruktuuri abil.

Ühte elementi hoitakse järgmises struktuuris:

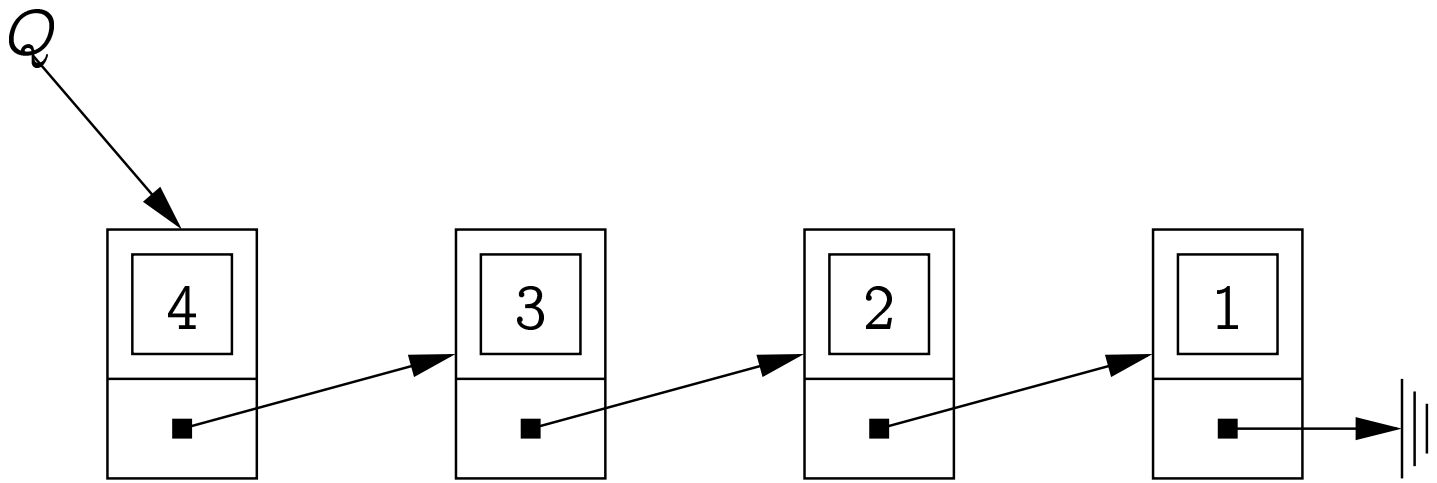


Siin *.järgm* on viit järgmisele (sügavamale) elemendile.

Magasini poole pöördumiseks kasutatakse viita tema esimesele elemendile.

Tühja magasini tähistab nullviit.

Magasin Q, millesse on lisatud elemendid $1, 2, 3, 4$:



$Q \Leftarrow r$ on siis

1 $e := \mathbf{new}(\text{magasini element})$
2 $e.\text{Andmed} := r$
3 $e.järgm := Q$
4 $Q := e$

ja $r \Leftarrow Q$ on

1 **if** $Q = \text{NIL}$ **then error**
2 $r := Q.\text{Andmed}$
3 $e := Q$
4 $Q := Q.järgm$
5 vabasta e

Järjekorda võib realiseerida dünaamilise andmestruktuuri abil.

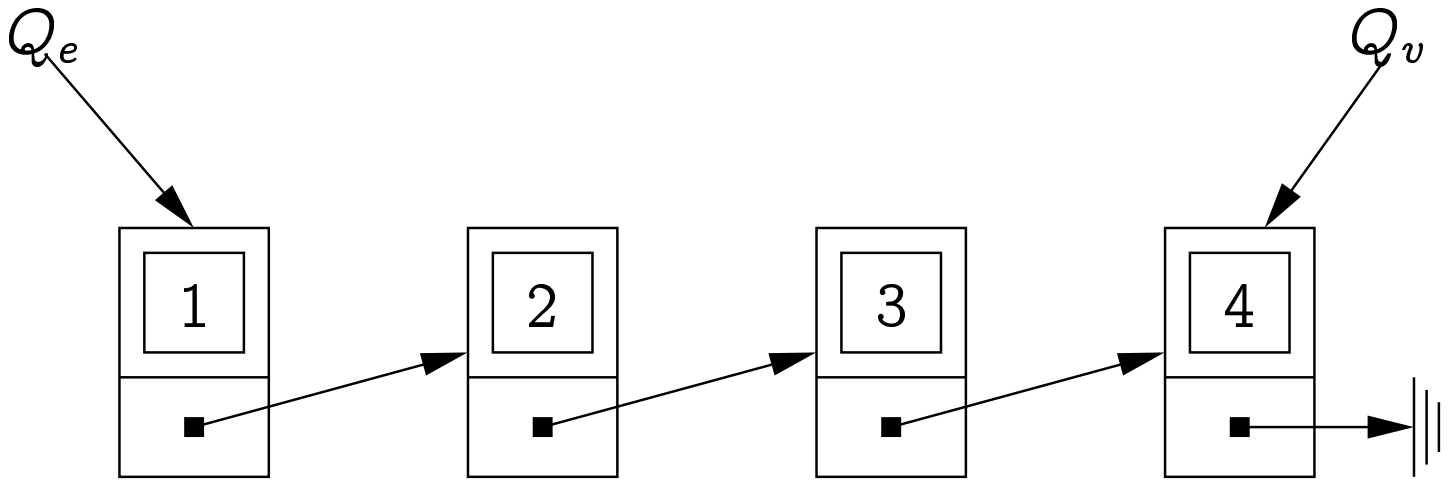
Ühte elementi hoitakse järgmises struktuuris:

Siin *.järgm* on viit järgmisele (hiljem lisatud) elemendile.

Järjekorra poole pöördumiseks kasutatakse viitasid tema esimesele ja viimasele elemendile.

Tühja järjekorda tähistab kaks nullviita.

Järjekord Q, kuhu on lisatud elemendid $1, 2, 3, 4$:



Siin Q_e on võtmise ja Q_v lisamise koht.

$Q \Leftarrow r$ on siis

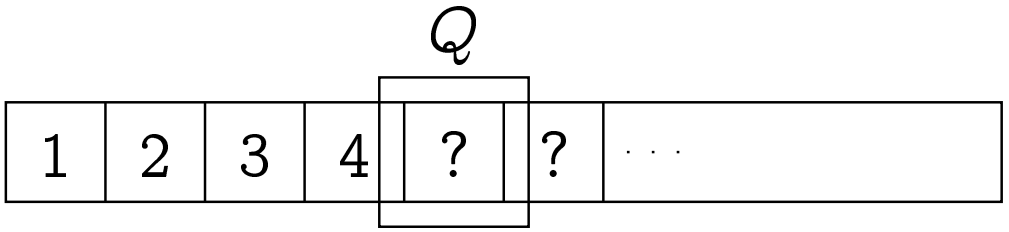
```
1   e := new(järjekorra element)
2   e.Andmed := r
3   e.järgm := NIL
4   if Q_e = Q_v = NIL then
5       Q_e := Q_v := e
6   else
7       Q_v.järgm := e
8   Q_v := e
```

ja $r \Leftarrow Q$ on

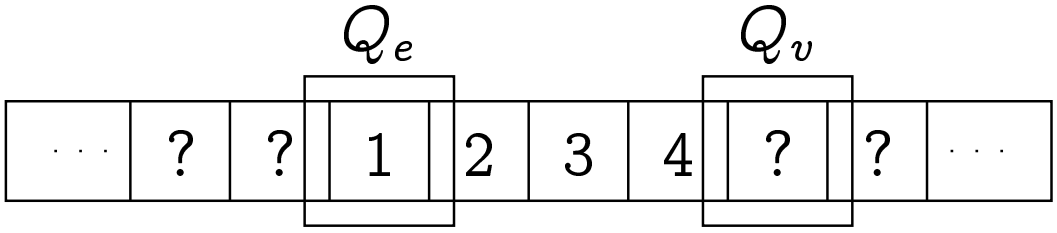
```
1   if Q_e = Q_v = NIL then error
2   r := Q_e.Andmed
3   e := Q_e
4   if Q_e = Q_v then
5       Q_e := Q_v := NIL
6   else
7       Q_e := Q_e.järgm
8   vabasta e
```

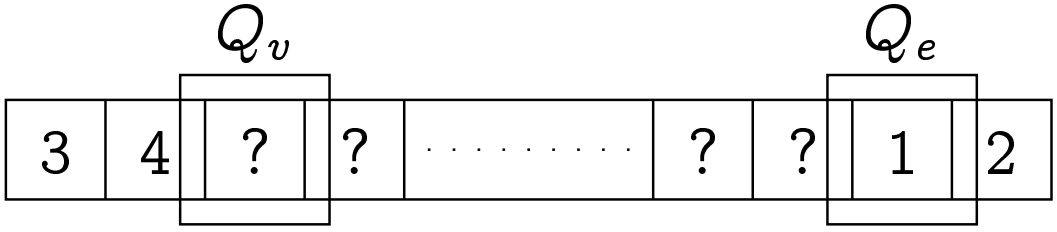
Kui on teada elementide suurim võimalik arv magasinis / järjekorras, siis saab teda realiseerida ka massiivi $a = [a_1, \ldots, a_n]$ abil.

Magasin, kuhu on lisatud elemendid $1, 2, 3, 4$:



Järjekord (2 varianti), kus on elemendid $1, 2, 3, 4$:





Olgu antud protseduur P, mida tahame välja kutsuda antud graafi G kõigil tippudel.

Kaks võimalikku järjekorda, mis mingis mõttes vastavad graafi struktuurile, on graafi *laiuti läbimine* ja *sügavuti läbimine*.

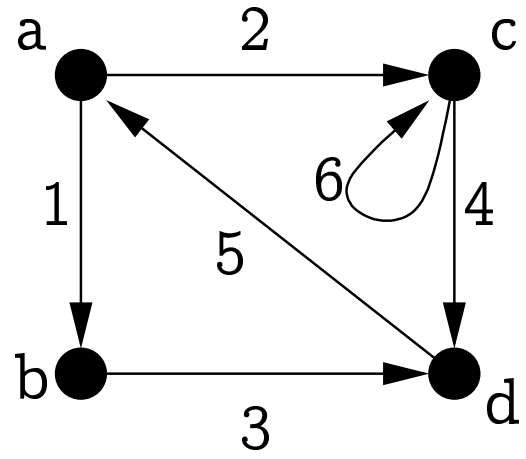
Olgu p viit graafi G esimesele tipule. Eeldame, et graafi kõik tipud on tipust p alates mööda servi liikudes saavutatavad.

Olgu tipu-/servastruktuuris täiendav väli *.läbitud* (ei kuulu ossa „Andmed“).

Laiuti läbimine: Olgu Q järjekord (esialgu tühi)

```
 1  q := p;
 2  while q ≠ NIL do
 3      q.läbitud := false; q := q.tipp
 4  p.läbitud := true; Q ⇐ p
 5  while ¬tühi?(Q) do
 6      q ⇐ Q; P(q);
 7      r := q.kaar
 8      while r ≠ NIL do
 9          if ¬(r.tipp.läbitud) then
10              Q ⇐ r.tipp
11              r.tipp.läbitud := true
12          r := r.kaar
```

Näide:



r

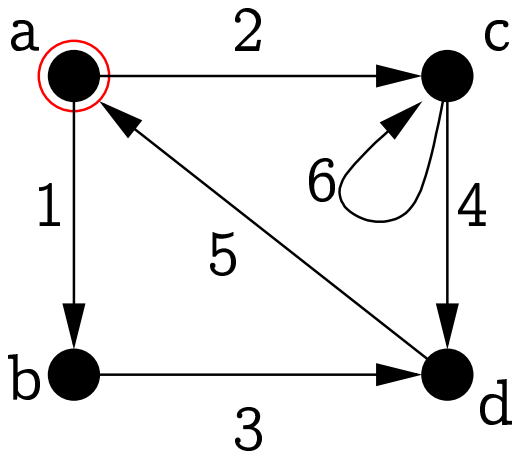
Q :

q

$P(\cdot)$ väljakutsed:

Näide:

r



q

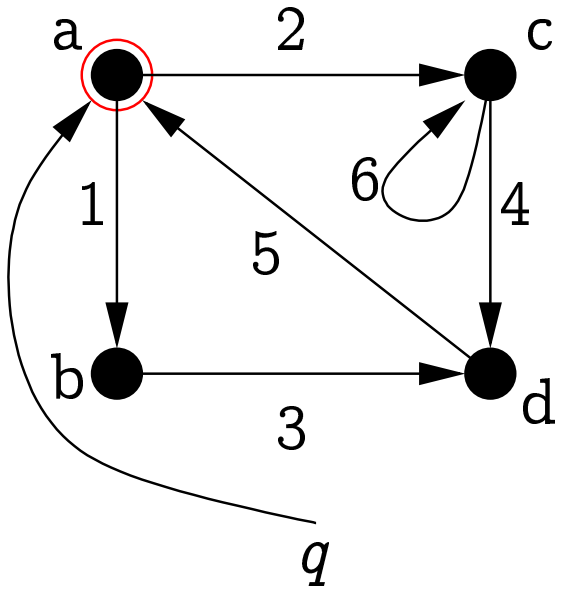
Q :

a

$P(\cdot)$ väljakutsed:

Näide:

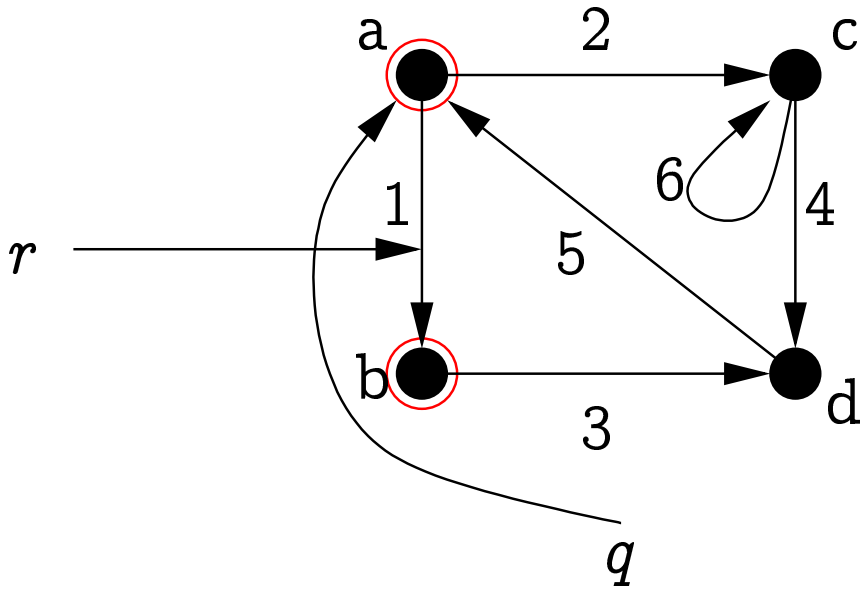
r



Q :

$P(\cdot)$ väljakutsed: a

Näide:

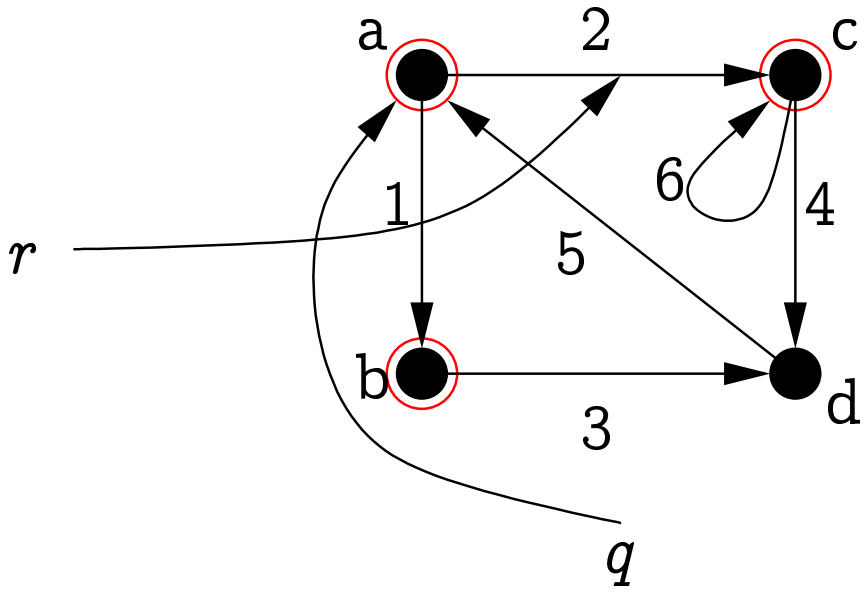


Q :

b

$P(\cdot)$ väljakutsed: a

Näide:

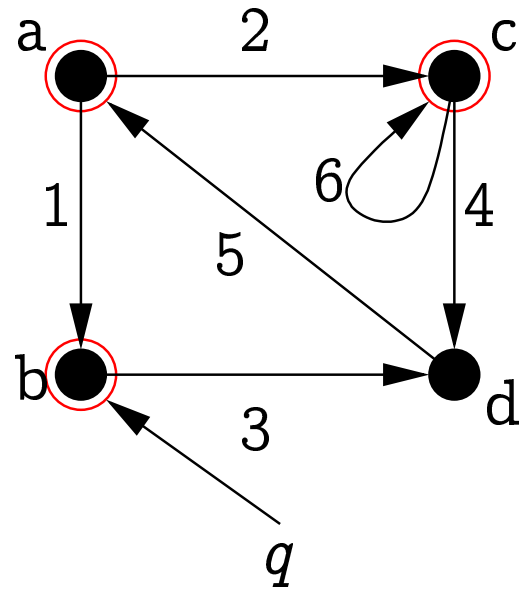


Q :

b c

$P(\cdot)$ väljakutsed: a

Näide:



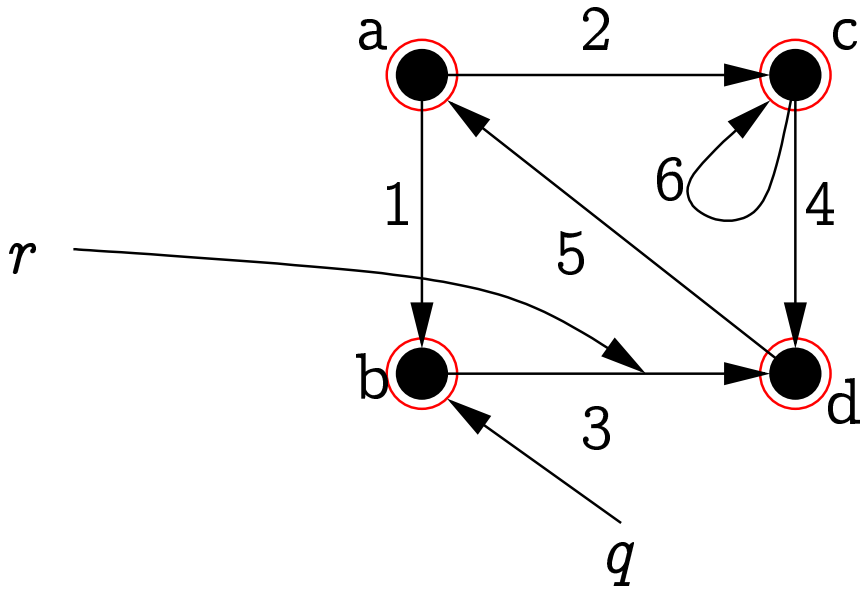
r

Q :

c

$P(\cdot)$ väljakutsed: a b

Näide:

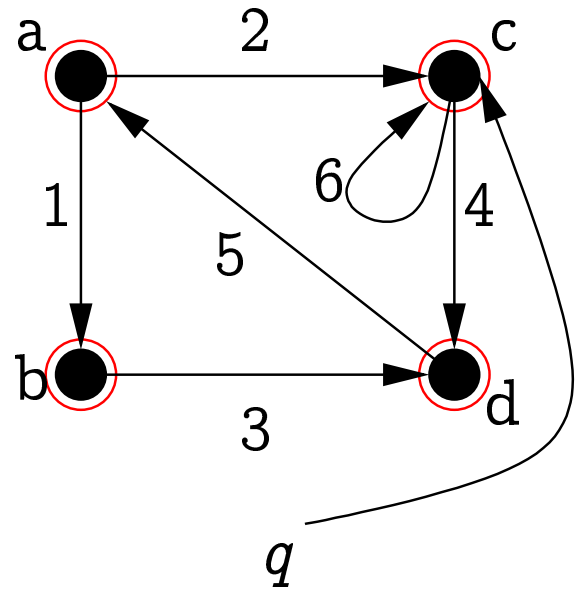


Q :

c d

$P(\cdot)$ väljakutsed: a b

Näide:



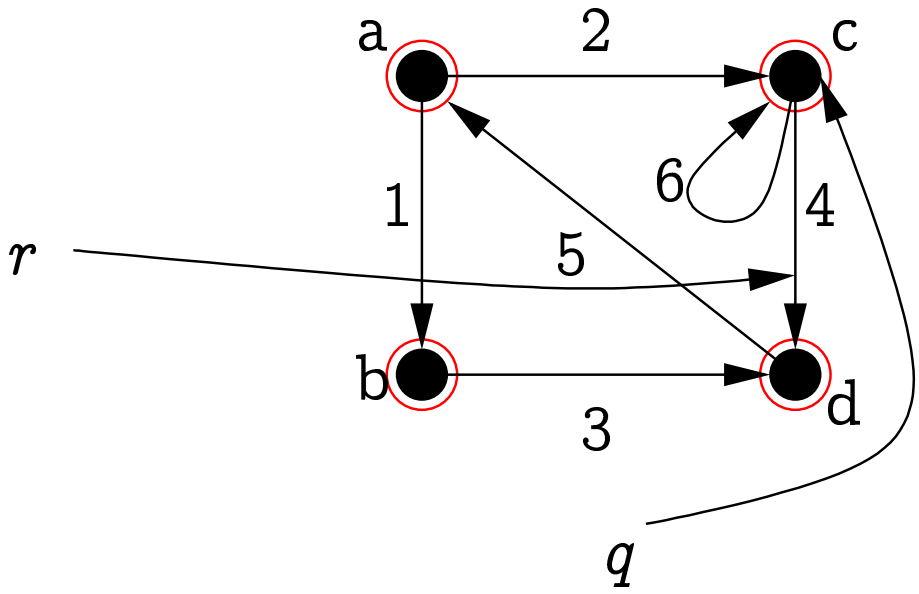
r

Q :

d

$P(\cdot)$ väljakutsed: a b c

Näide:

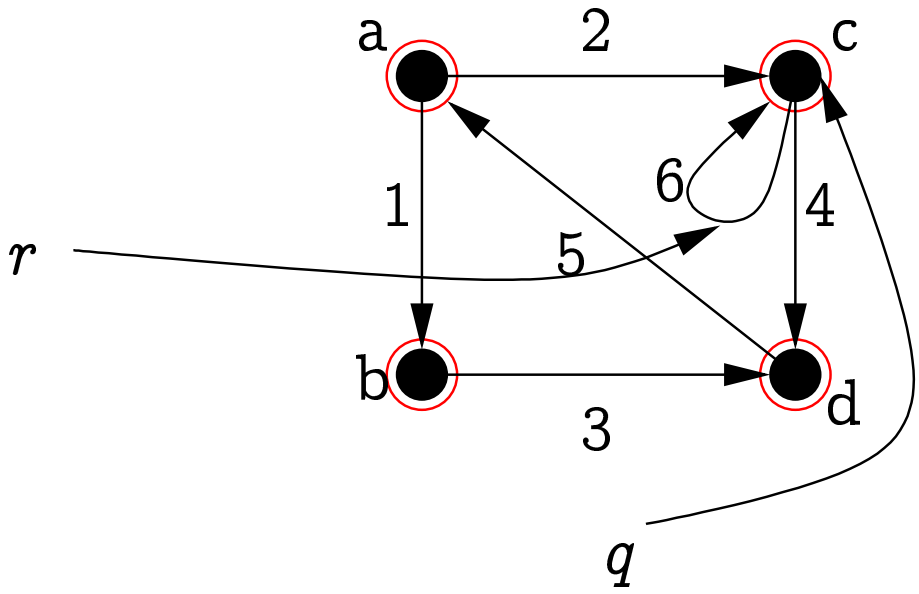


Q :

d

$P(\cdot)$ väljakutsed: a b c

Näide:



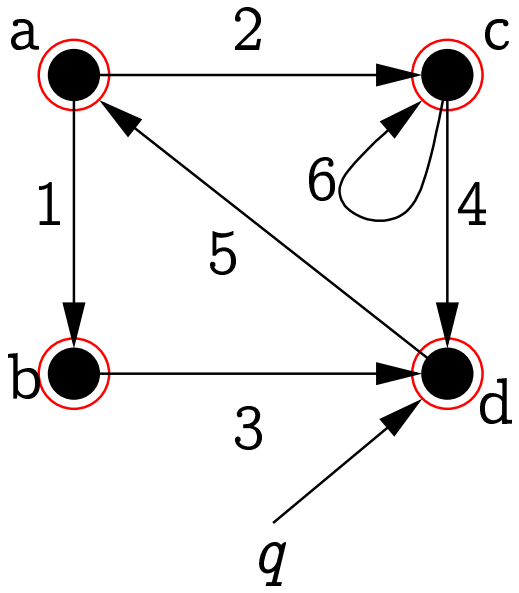
Q :

d

$P(\cdot)$ väljakutsed: a b c

Näide:

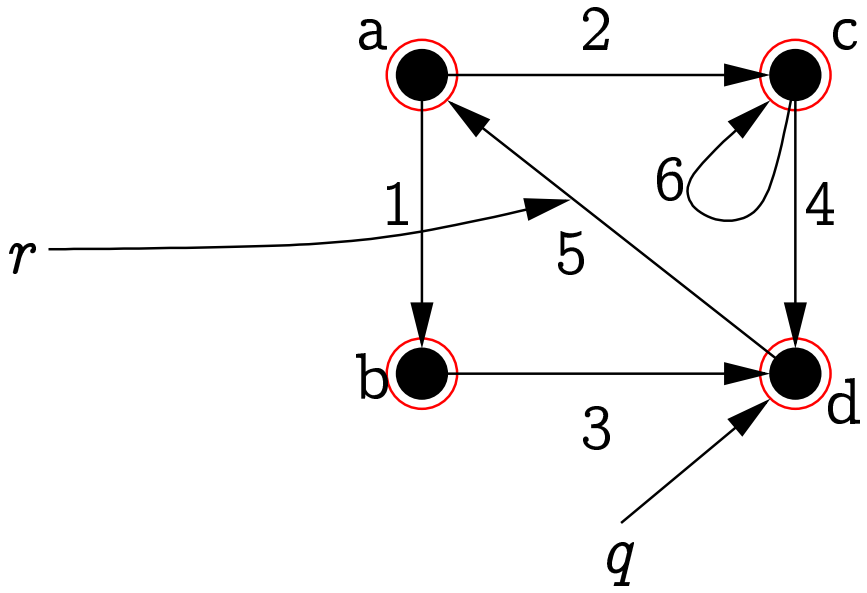
r



Q :

$P(\cdot)$ väljakutsed: a b c d

Näide:

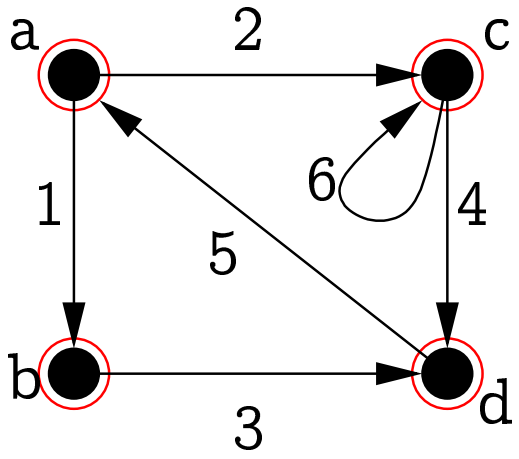


Q :

$P(\cdot)$ väljakutsed: a b c d

Näide:

r



q

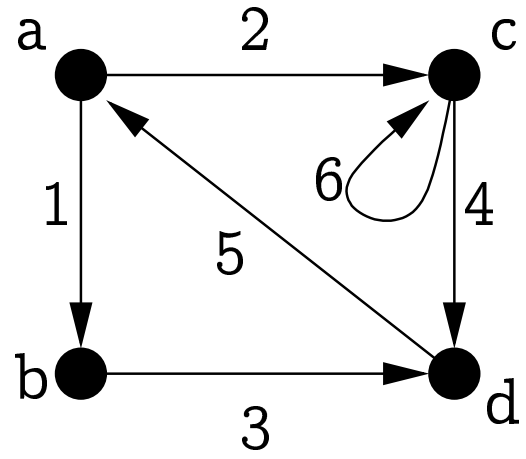
Q :

$P(\cdot)$ väljakutsed: a b c d

Sügavuti läbimine: Olgu Q magasin (esialgu tühi)

1..12 sama programm, mis laiuti läbimisel

Näide:



r

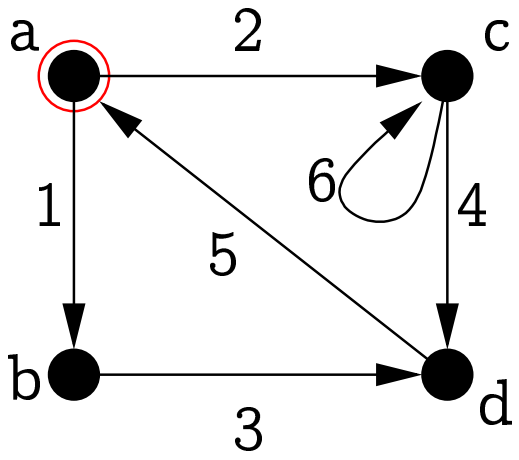
Q :

q

$P(\cdot)$ väljakutsed:

Näide:

r



q

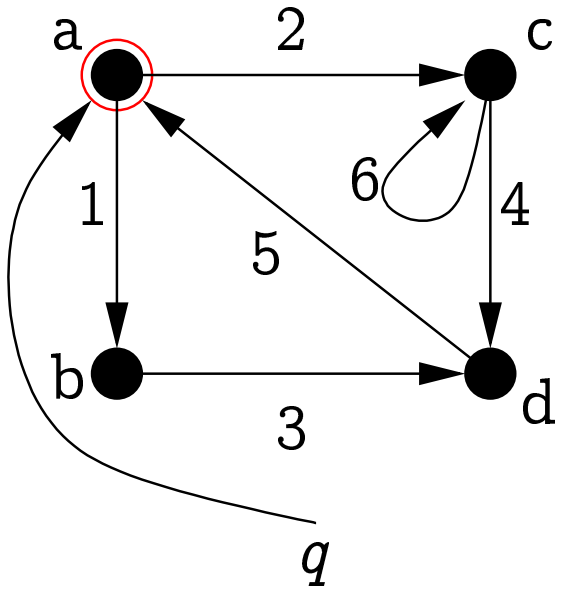
Q :

a

$P(\cdot)$ väljakutsed:

Näide:

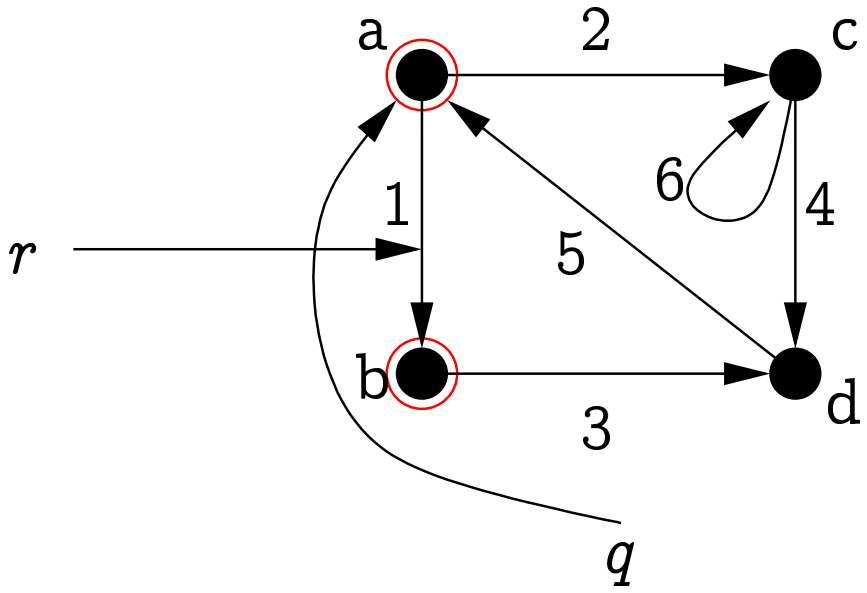
r



Q :

$P(\cdot)$ väljakutsed: a

Näide:

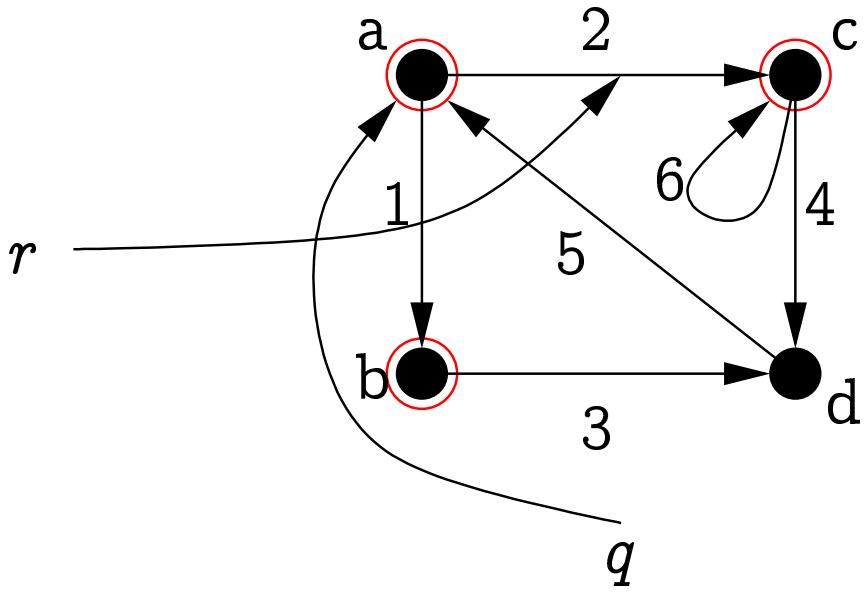


Q :

b

$P(\cdot)$ väljakutsed: a

Näide:

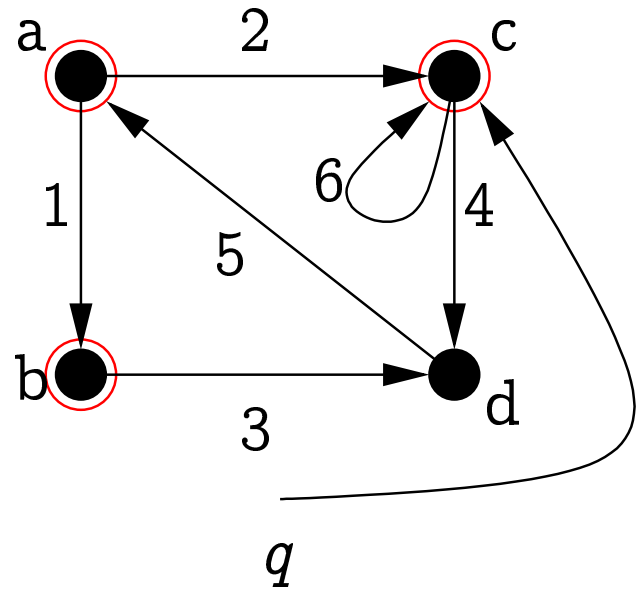


Q :

b c

$P(\cdot)$ väljakutsed: a

Näide:



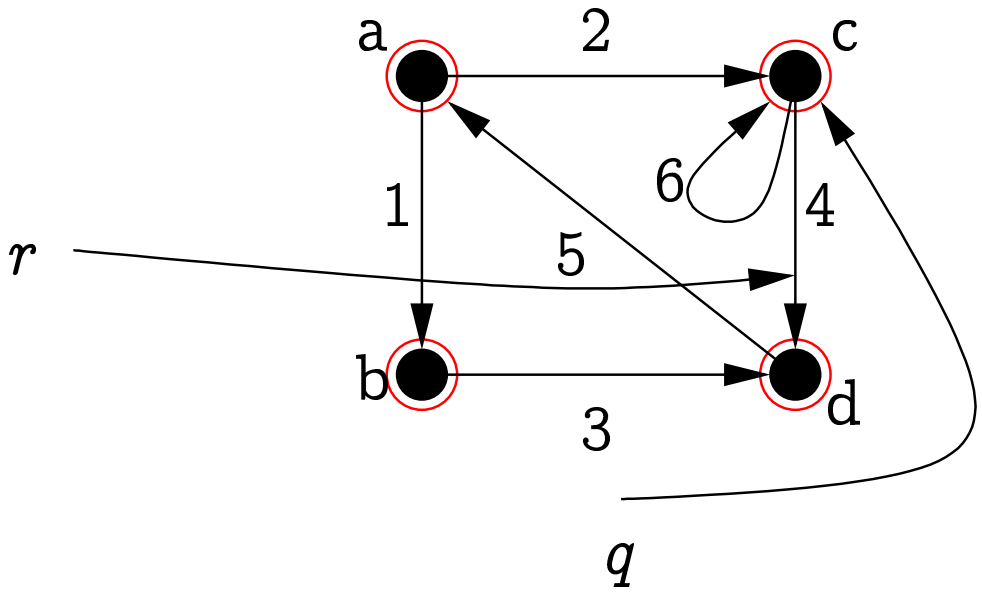
r

Q :

b

$P(\cdot)$ väljakutsed: a c

Näide:

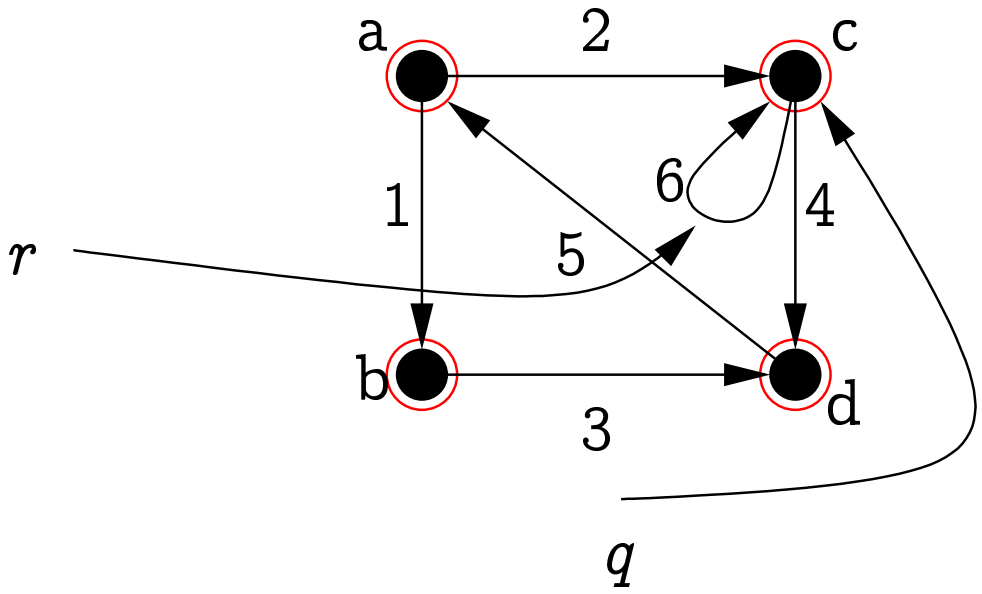


Q :

b d

$P(\cdot)$ väljakutsed: a c

Näide:

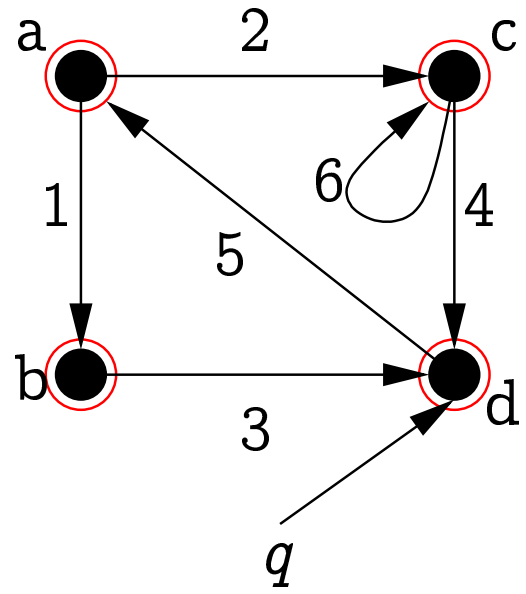


Q :

b d

$P(\cdot)$ väljakutsed: a c

Näide:



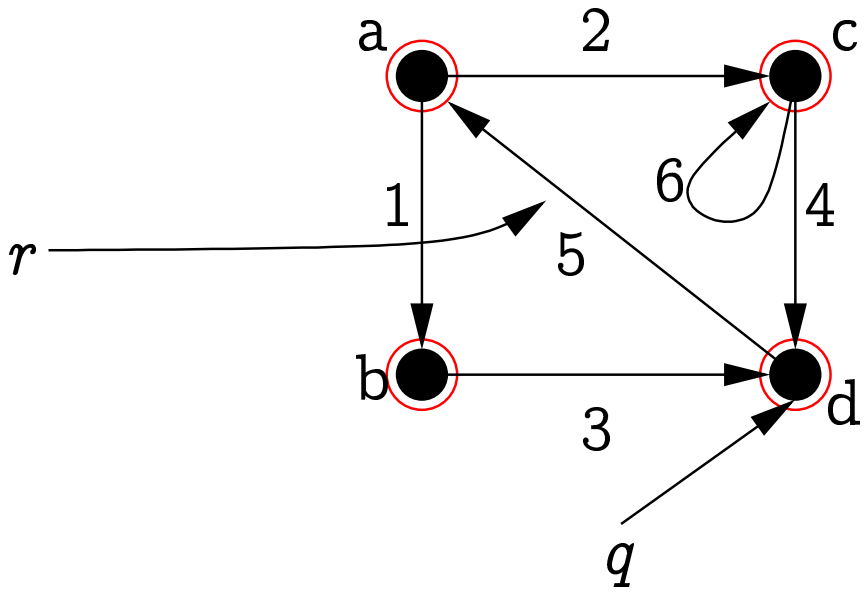
r

Q :

b

$P(\cdot)$ väljakutsed: a c d

Näide:



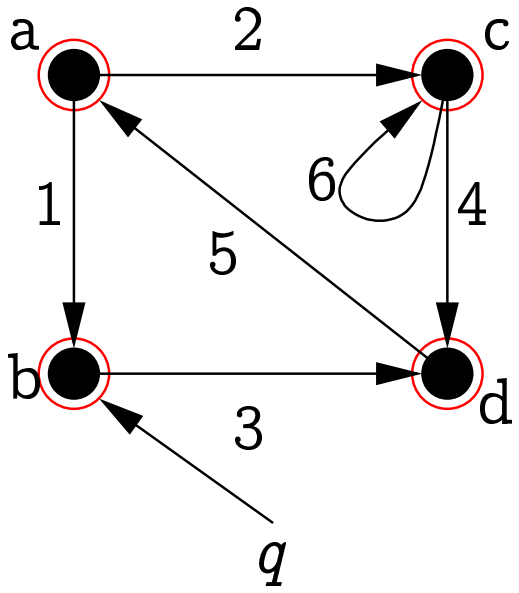
Q :

b

$P(\cdot)$ väljakutsed: a c d

Näide:

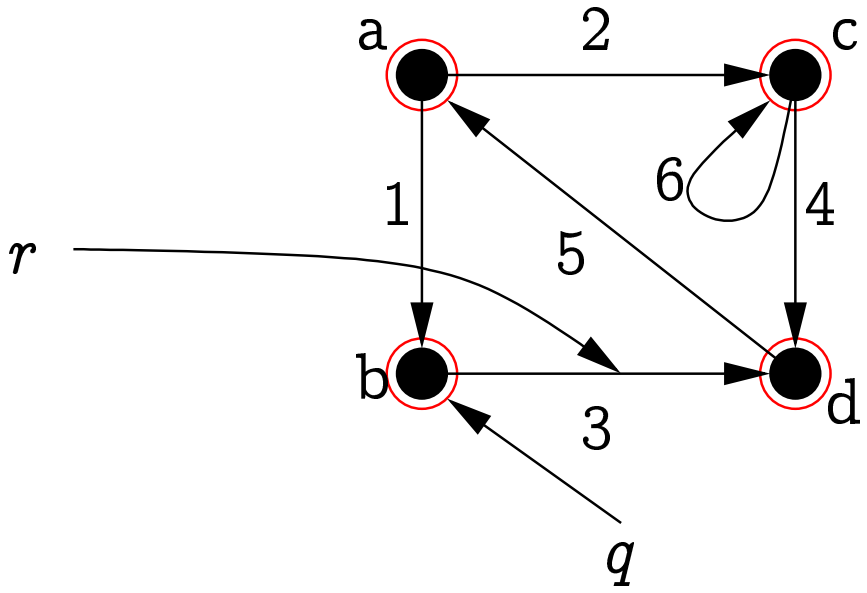
r



Q :

$P(\cdot)$ väljakutsed: a c d b

Näide:

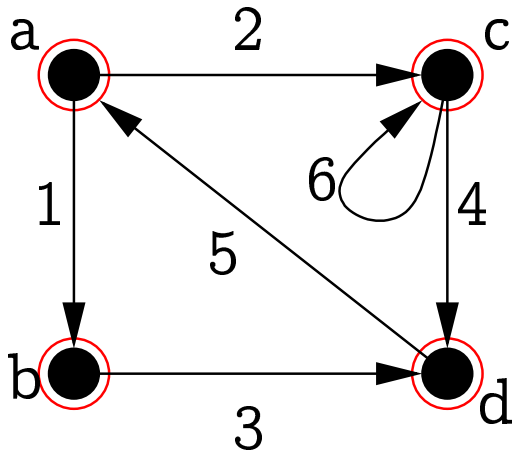


Q :

$P(\cdot)$ väljakutsed: a c d b

Näide:

r



q

Q :

$P(\cdot)$ väljakutsed: a c d b

Antud suunatud graaf $G = (V, E)$, iga serva $e \in E$ jaoks defineeritud tema *pikkus* $\ell(e)$.

Tee $v_0 \xrightarrow{e_1} v_1 \xrightarrow{e_2} v_2 \xrightarrow{e_3} \cdots \xrightarrow{e_n} v_n$ *pikkus* on $\ell(e_1) + \ell(e_2) + \cdots + \ell(e_n)$.

Antud kaks tippu $u, v \in V$. Leida nendevaheline lühim tee.

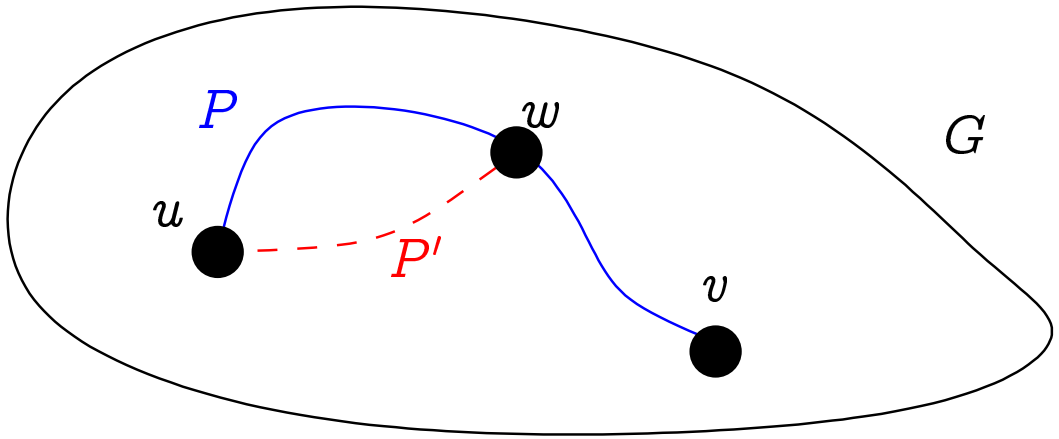
Tähistusi:

- Kui $v \in V$, siis Gv tähistagu kõigi selliste tippude w hulka, kuhu on tipust v serv ning $G^{-1}v$ kõigi selliste tippude w hulka, kust on tippu v serv.
- Kui $u, v \in V$, siis olgu $\ell(u, v)$ lühima u-st v-sse mineva serva pikkus. Kui u-st v-sse ei lähe ühtegi serva, siis olgu $\ell(u, v) = \infty$.
- Kui $u, v \in V$, siis olgu $d(u, v)$ lühima tee pikkus u-st v-sse. Kui u-st ei saa v-sse, siis $d(u, v) = \infty$.
- Kui $u, v \in V$ ja $i \in \mathbb{N}$, siis olgu $D_i(u, v)$ sellise lühima tee pikkus u-st v-sse, kus on ülimalt i serva. Kui selliseid teid ei ole, siis $D_i(u, v) = \infty$.

Ülesandel on optimaalne alamstruktuur:

Olgu P lühim tee tipust u tipuni v. Olgu w mingi tipp sellel teel. Siis tee P algusosa tipust u tipuni w on lühim tee u-st w-sse.

Tõepoolest, kui leiduks lühem tee P' u-st w-sse, siis võiks P algusosa asendada P'-ga ja saada P-st lühema tee u-st v-sse.



Järeldus: iga u, v, w jaoks $d(u,v) \leqslant d(u,w) + \ell(w,v)$.

Kehtib $D_0(u,v) \geqslant D_1(u,v) \geqslant D_2(u,v) \geqslant \cdots$. Peale selle, mingi i jaoks $D_i(u,v) = D_{i+1}(u,v) = \cdots = d(u,v)$.

Selline i leidub küll ainult siis, kui graafis pole *negatiivse pikkusega tsükleid*. Siis $i \leqslant |V| - 1$, sest lühim tee läbib iga tippu ülimalt ühel korral.

Kehtib (alati)

$$D_{i+1}(u,v) = \min\big(\{D_i(u,v)\} \cup \{\, D_i(u,w) + \ell(w,v) : w \in G^{-1}v\}\big),$$

seega saab negatiivse pikkusega tsüklite mitteesinemisel leida tippude kaugusi tipust u järgmisel viisil:

Olgu D ja DD massiivid, mis on indekseeritud graafi $G = (V, E)$ tippudega.

```
 1  D[u] := 0
 2  for all v ∈ V\{u} do D[v] := ∞
 3  for all v ∈ Gu do D[v] := ℓ(u, v)
 4  repeat
 5      DD := D
 6      for all v ∈ V do
 7          d := DD[v]
 8          for all w ∈ G^{-1}v do
 9              d := min(d, DD[w] + ℓ(w, v))
10          D[v] := d
11  until D = DD
12  return D
```

Invariant: peale i-ndat iteratsiooni (reas 11) on $D[v] = D_i(u, v)$.

Ridades 7–9 toimuvat miinimumi leidmist võib ka otse väljal $D[v]$ teha, s.t. abimuutujat d kasutamata. Siis omistamist $D[v] := DD[v]$ pole vaja teha, sest need on niikuinii võrdsed (5. rea tõttu).

```
 1  D[u] := 0
 2  for all v ∈ V\{u} do D[v] := ∞
 3  for all v ∈ Gu do D[v] := ℓ(u, v)
 4  repeat
 5      DD := D
 6      for all v ∈ V do
 7          for all w ∈ G^{-1}v do
 8              D[v] := min(D[v], DD[w] + ℓ(w, v))
 9  until D = DD
10  return D
```

Siin 6. ja 7. rida korraldavad lihtsalt tsükli, mis graafi kõik servad korra läbi käib.

Ta on korraldatud nii, et käime kõigepealt läbi servade võimalikud lõpptipud v ja seejärel kõik servad, mis selles tipus lõppevad.

Tsüklit üle servade võib ka teisiti korraldada, näiteks käia kõigepealt läbi kõik võimalikud algtipud ja seejärel kõik servad, mis sellest tipust algavad.

Levinud graafiesitusviiside korral (tippude lihtahel ja iga tipu juures temast algavate servade lihtahel) võib selline korraldus efektiivsem olla.

```
 1  D[u] := 0
 2  for all v ∈ V\{u} do D[v] := ∞
 3  for all v ∈ Gu do D[v] := ℓ(u, v)
 4  repeat
 5      DD := D
 6      for all w ∈ V do
 7          for all v ∈ Gw do
 8              D[v] := min(D[v], DD[w] + ℓ(w, v))
 9  until D = DD
10  return D
```

Kui meil on negatiivse pikkusega tsükleid, siis sellist i-d ei leidu, et $D_i(u,v) = D_{i+1}(u,v) = D_{i+2}(u,v) = \cdots$.

Siis pole isegi $d(u,v)$ defineeritud.

Kuidas negatiivse pikkusega tsüklite olemasolu avastada?

Kui ei ole negatiivse pikkusega tsükleid, siis on iteratsioone, kus midagi muutub, ülimalt $|V| - 1$.

Teemegi ülimalt $|V|$ iteratsiooni ja kui viimasel iteratsioonil midagi muutus, siis anname veateate.

```
 1  D[u] := 0
 2  for all v ∈ V\{u} do D[v] := ∞
 3  for all v ∈ Gu do D[v] := ℓ(u, v)
 4  i := 0
 5  repeat
 6      DD := D
 7      for all w ∈ V do
 8          for all v ∈ Gw do
 9              D[v] := min(D[v], DD[w] + ℓ(w, v))
10      i := i + 1
11  until D = DD or i = |V|
12  if D = DD then return D
13  error „negatiivse pikkusega tsükkel“
```

Hakkama saab ka ilma massiivita DD:

```
 1  D[u] := 0
 2  for all v ∈ V\{u} do D[v] := ∞
 3  for all v ∈ Gu do D[v] := ℓ(u, v)
 4  i := 0
 5  repeat
 6      muutus := false
 7      for all w ∈ V do
 8          for all v ∈ Gw do
 9              if D[v] > D[w] + ℓ(w, v) then
10                  D[v] := D[w] + ℓ(w, v)
11                  muutus := true
12      i := i + 1
13  until ¬muutus or i = |V|
14  if ¬muutus then return D
15  error „negatiivse pikkusega tsükkel“
```

Tsükliinvariant on antud juhul (tsükli lõpus, real 13):

Peale i-ndat iteratsiooni on $D[v]$ minimaalse pikkusega tee mingite u-st v-sse viivate teede hulgas. Seejuures see hulk sisaldab kõiki teid u-st v-sse, milles on ülimalt i serva.

Peale eelviimast iteratsiooni sisaldab see hulk seega kõiki teid u-st v-sse, mis läbivad ülimalt $(|V| - 1)$-t serva. Muu-hulgas siis ka lühimat teed u-st v-sse.

Esitatud algoritm on *Bellman-Ford*i algoritm.

Keerukus:

Välimist tsüklit (read 5–13) itereeritakse kuni $|V|$ korda.

Selle sees on sisemine tsükkel (read 7–11) üle kõigi servade.
Selle tsükli keha (read 9–11) võtab konstantse aja.

Kokku seega $\Theta(|V| \cdot |E|)$.

Toodud algoritm leiab lühimate teede pikkused tipust u teistesse tippudesse.

Kuidas leida/salvestada teid ise?

Iga tipu v jaoks salvestame talle eelneva tipu (senileitud) lühimal teel u-st v-sse. Kasutame selleks massiivi π (indekseeritud graafi tippudega).

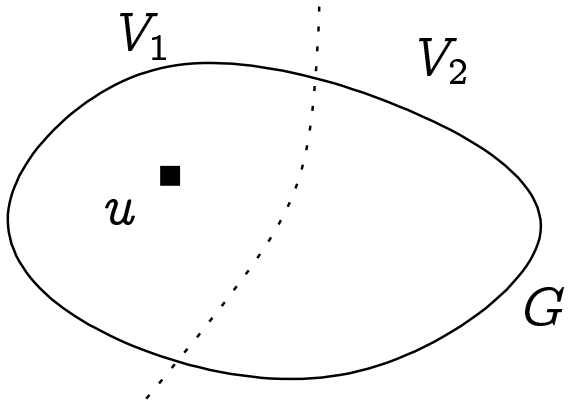
Lühim tee u-st v-sse on siis (ümberpööratult)
$v, \pi[v], \pi[\pi[v]], \pi[\pi[\pi[v]]], \ldots, u$.

Massiivid D ja π rahuldavad iga tipu v jaoks, kus $\pi[v]$ on defineeritud: $D[v] = D[\pi[v]] + \ell(\pi[v], v)$.

```
 1  D[u] := 0
 2  for all v ∈ V\{u} do D[v] := ∞
 3  for all v ∈ Gu do D[v] := ℓ(u, v); π[v] := u
 4  i := 0
 5  repeat
 6      muutus := false
 7      for all w ∈ V do
 8          for all v ∈ Gw do
 9              if D[v] > D[w] + ℓ(w, v) then
10                  D[v] := D[w] + ℓ(w, v)
11                  π[v] := w
12                  muutus := true
13      i := i + 1
14  until ¬muutus or i = |V|
15  if ¬muutus then return (D, π)
16  error „negatiivse pikkusega tsükkel“
```

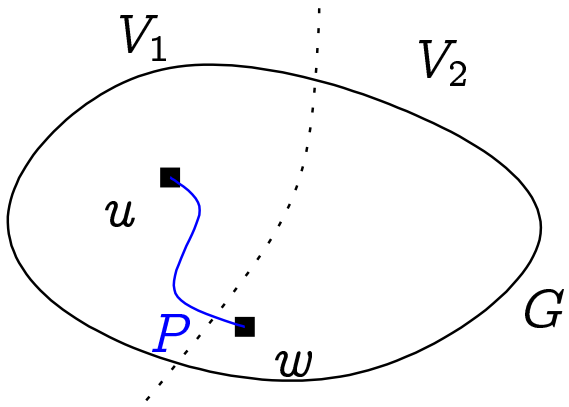
Eeldame nüüd, et negatiivse pikkusega servi ei leidu. Siis saab lühimaid teid u-st teistesse tippudesse kiiremini leida.

Oletame, et me oleme juba leidnud V_1-e kuuluvate tippude jaoks lühimate teede pikkused u-st nendesse tippudesse.



Samuti oletame, et ükski V_1-e kuuluv tipp pole u-st kaugemal kui ükski V_2-e kuuluv tipp.

Olgu w u-le lähim tipp osas V_2. Olgu P lühim tee u-st w-sse.



Siis P kõik tipud peale w asuvad osas V_1. Kui veel mõni P tippudest asuks osas V_2, siis oleks see tipp u-le lähemal kui w.

Tahame seda w-d leida. Me eemaldaks ta V_2-st ja lisaks V_1-e. Eelmise slaidi all toodud invariant jääks siis kehtima.

Leidmaks seda tippu w:

Me peame vaatama ainult selliseid ahelaid u-st V_2-e tippudesse, mille kõik tipud, peale viimase, asuvad V_1-s.

Olgu antud V_1 ja V_2, samuti olgu iga $v \in V_1$ jaoks antud $D[v]$ — lühima tee pikkus u-st v-sse.

Vaatame kõiki servi (v, w), kus $v \in V_1$ ja $w \in V_2$. Leiame $D[v] + \ell(v, w)$. Leiame v ja w, mille korral see summa minimaalne on.

Lisame w V_1-e ning võtame $D[w] := D[v] + \ell(v, w)$ ja $\pi[w] := v$.

Selle w kiiresti leidmiseks:

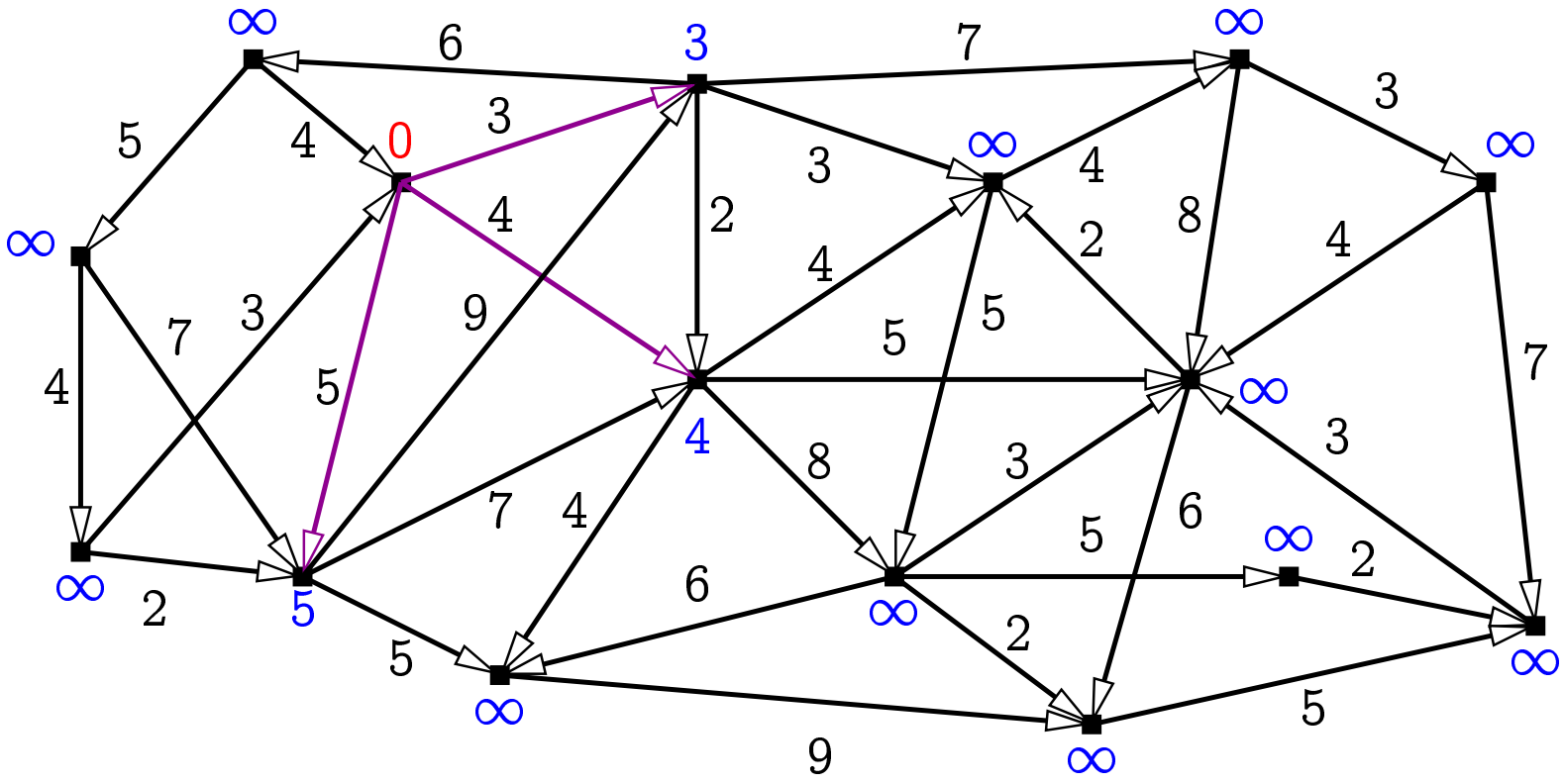
Iga $w \in V_2$ jaoks olgu tema kaal $\delta(w)$ defineeritud kui

$$\delta(w) := \min_{v \in V_1} D[v] + \ell(v, w) \ .$$

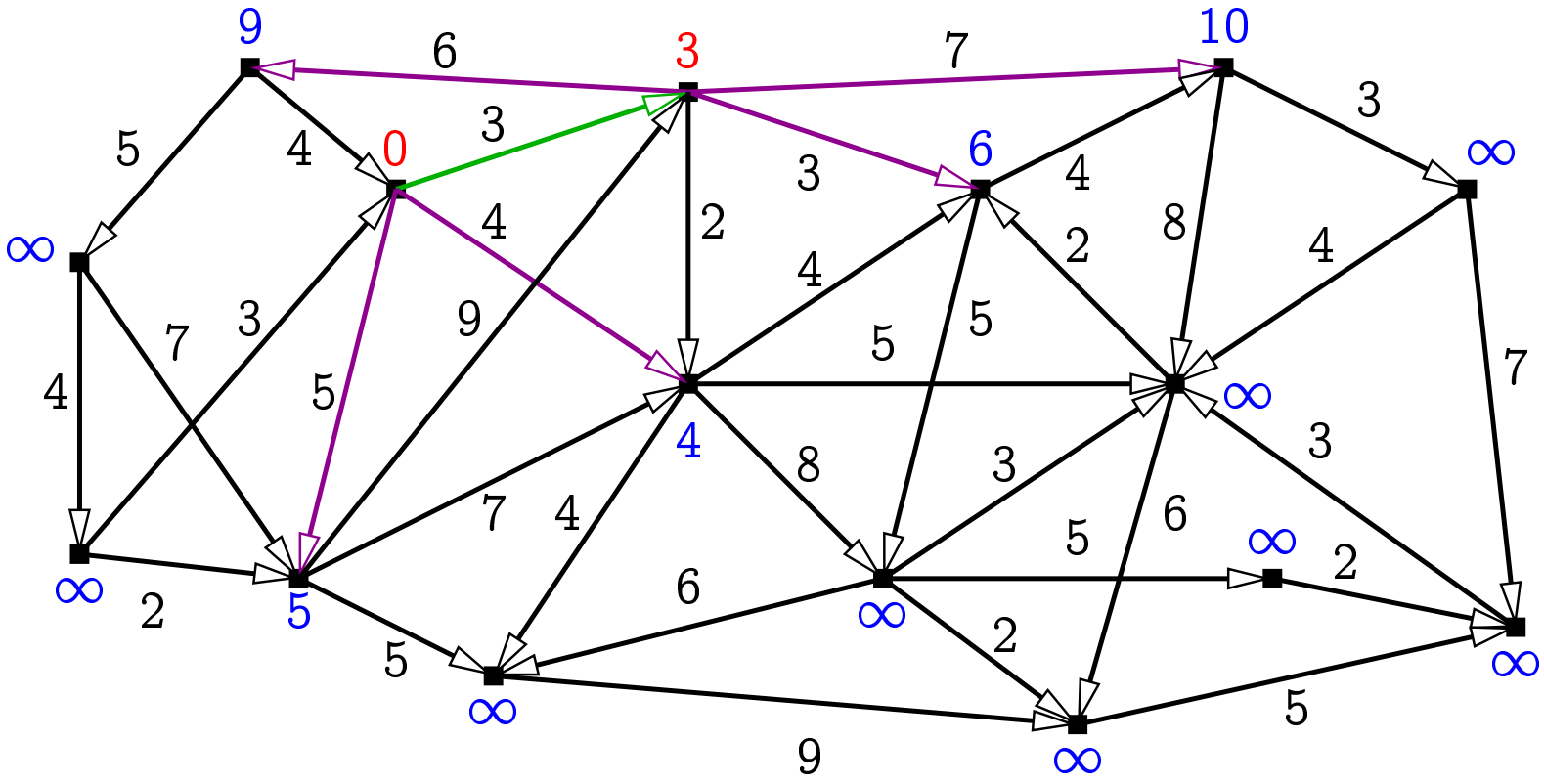
Hoiame V_2 elemente kuhjas (võtmeks on kaal $\delta(\cdot)$). Siis saab minimaalse $\delta(\cdot)$-ga tippu lihtsalt leida.

Kui me tipu w viime V_2-st V_1-e, siis võivad w naabertippude kaalud väheneda.

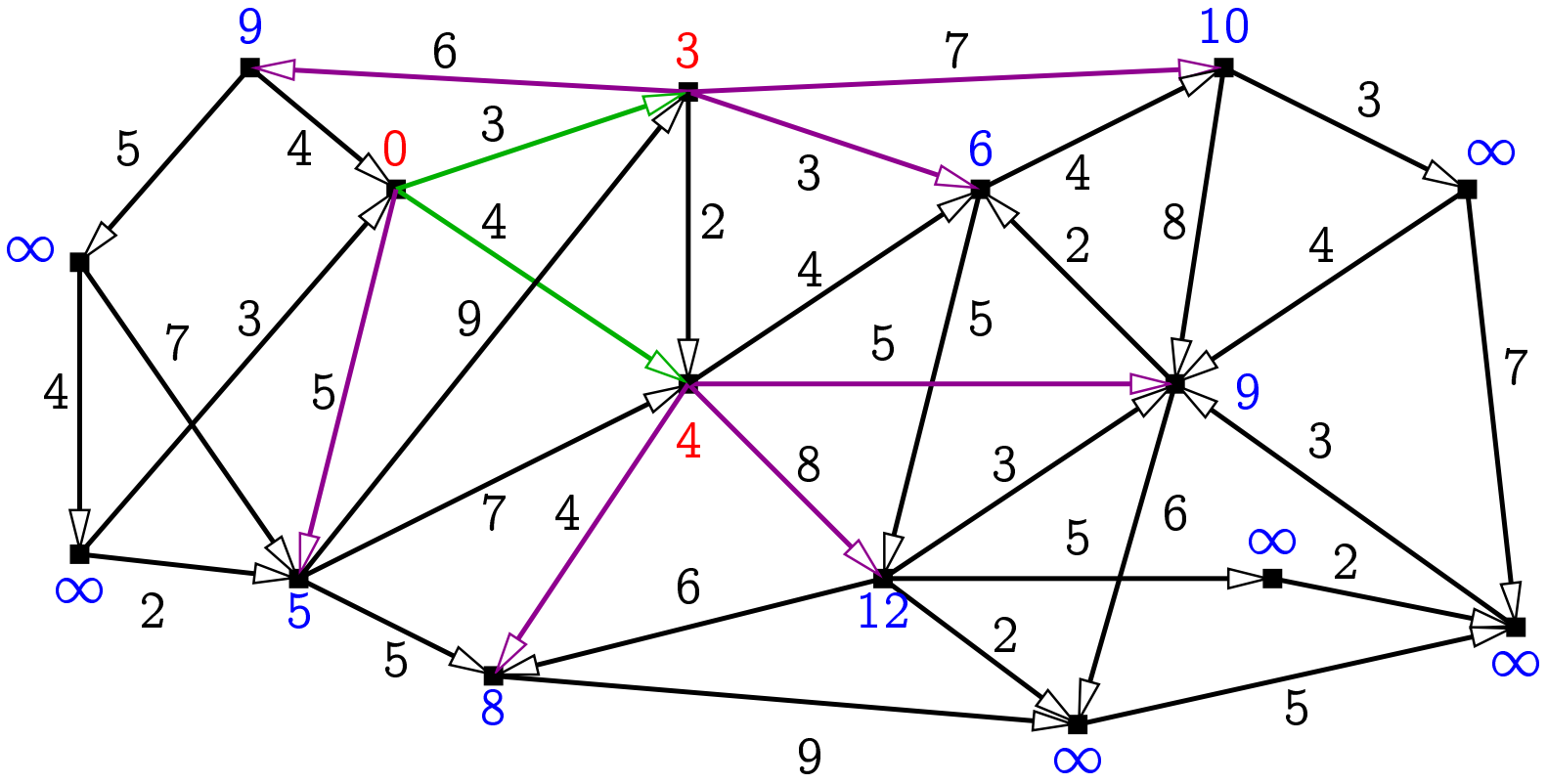
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



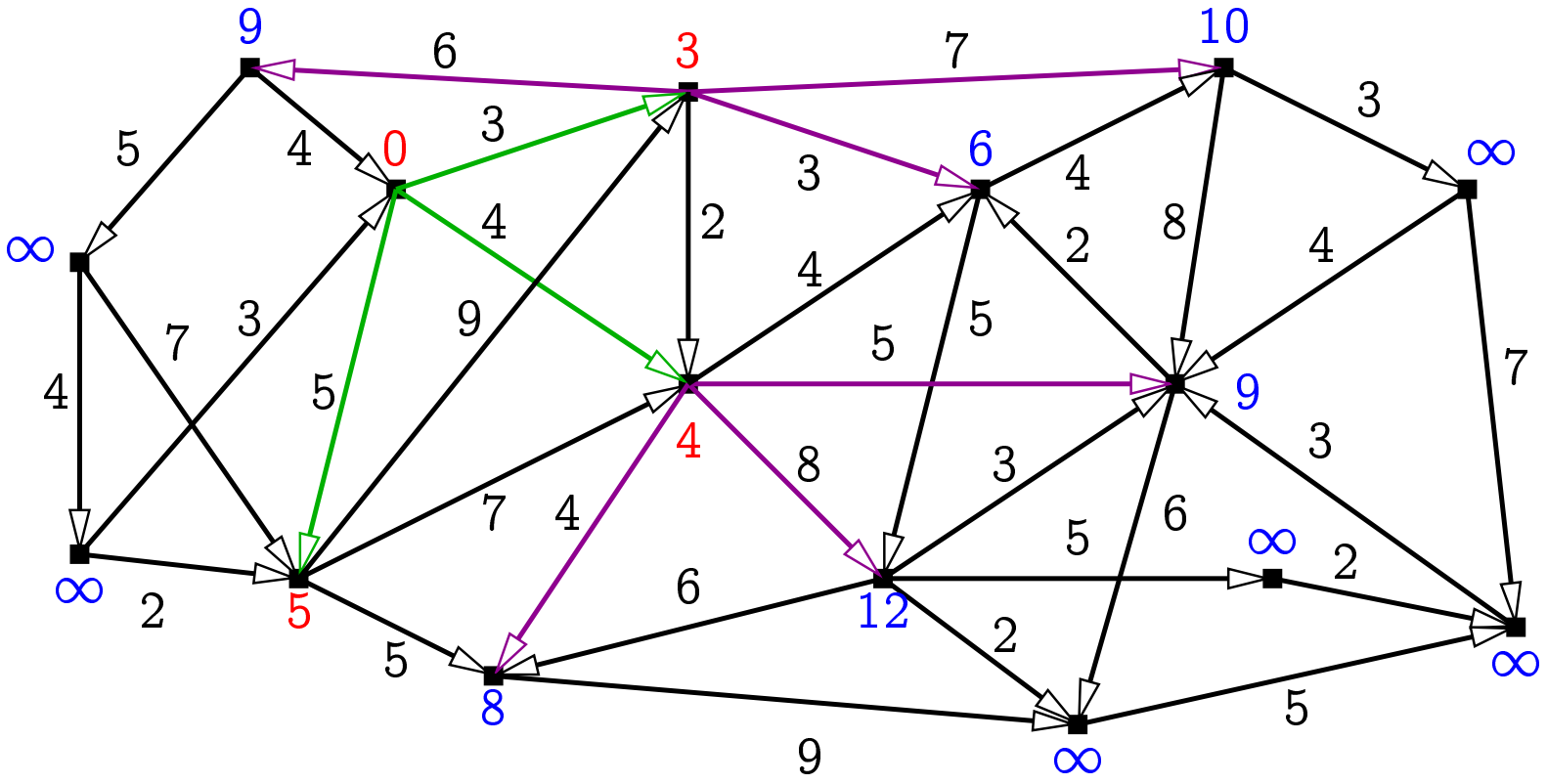
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



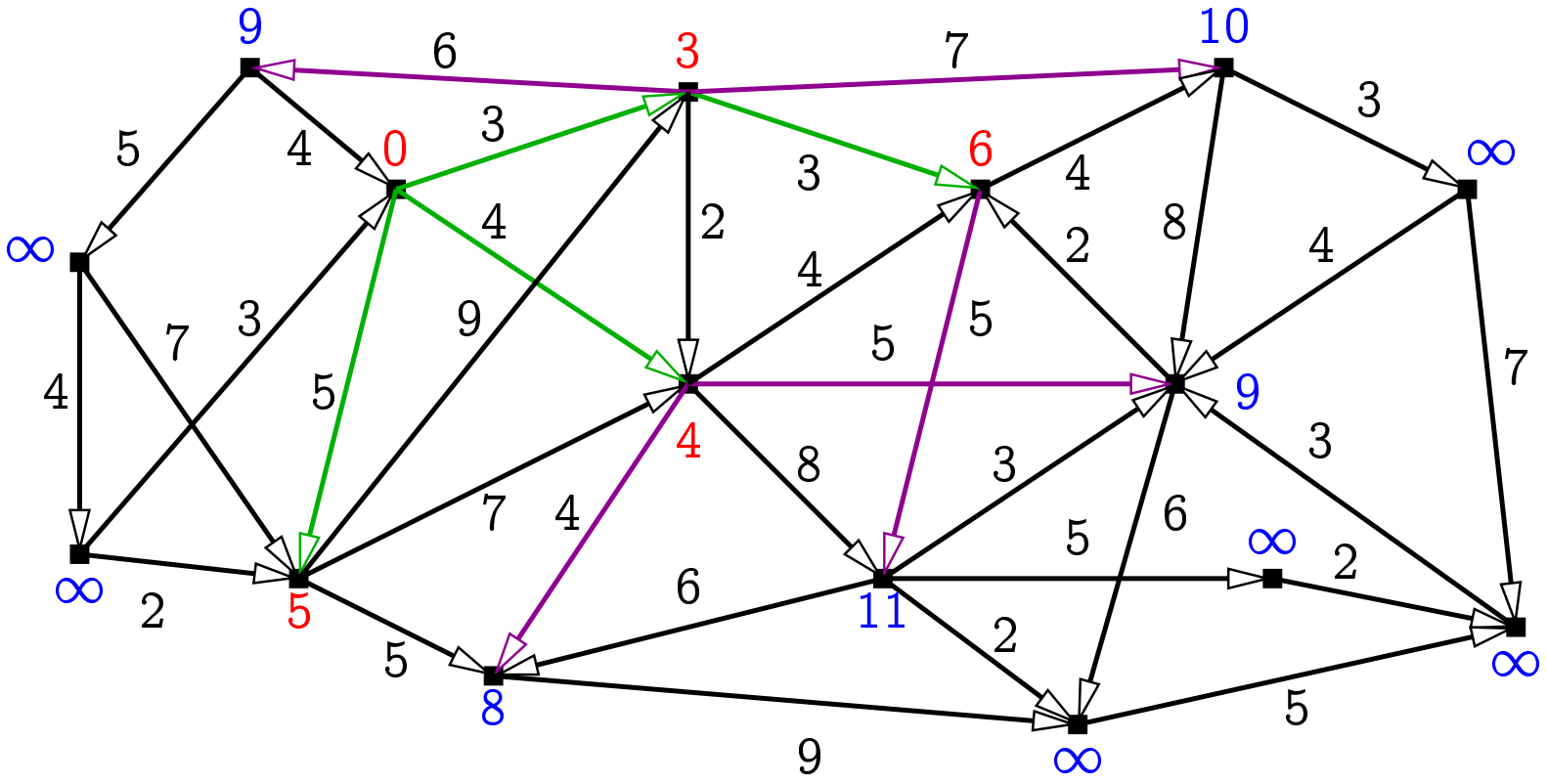
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



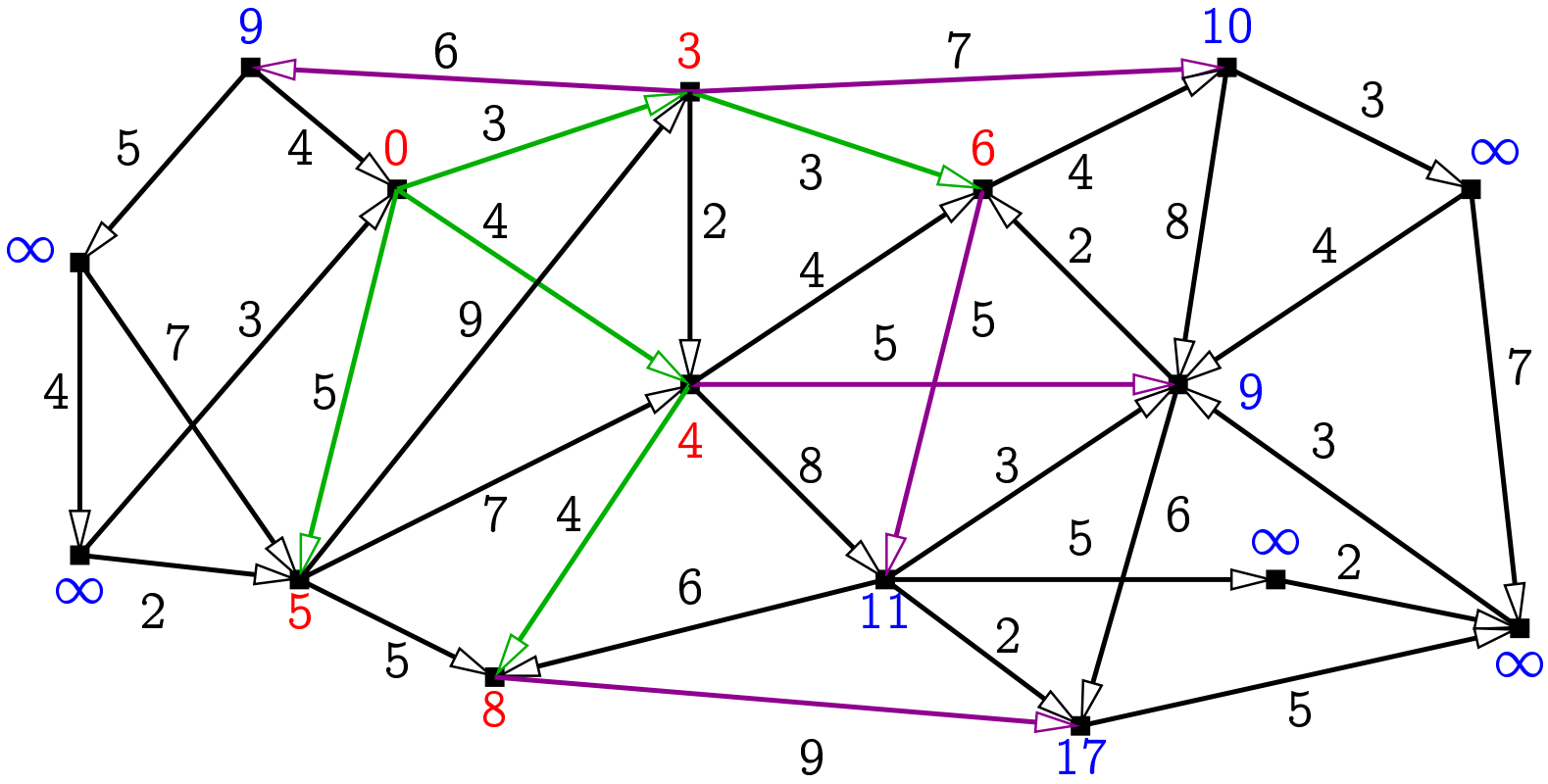
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



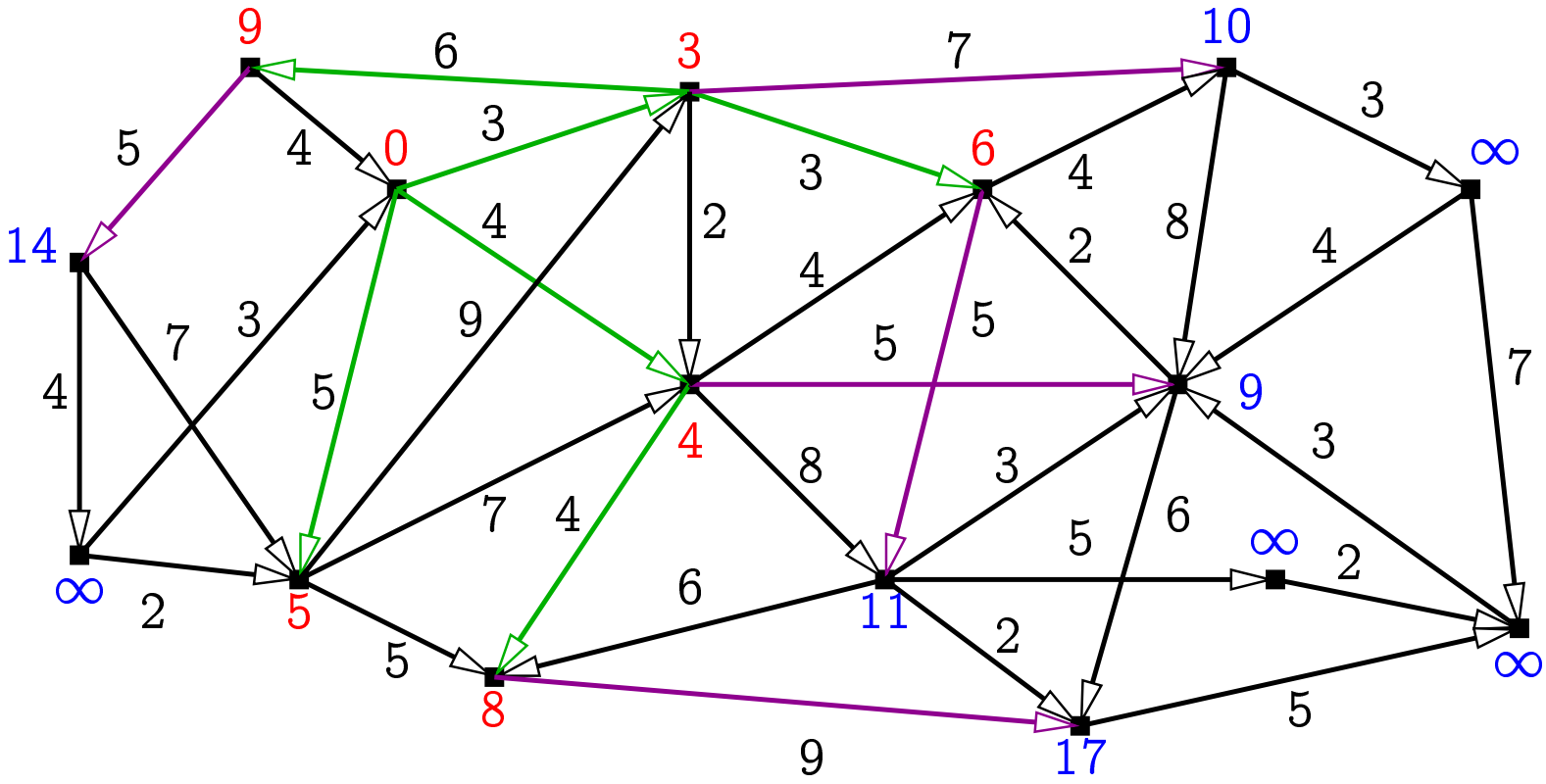
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



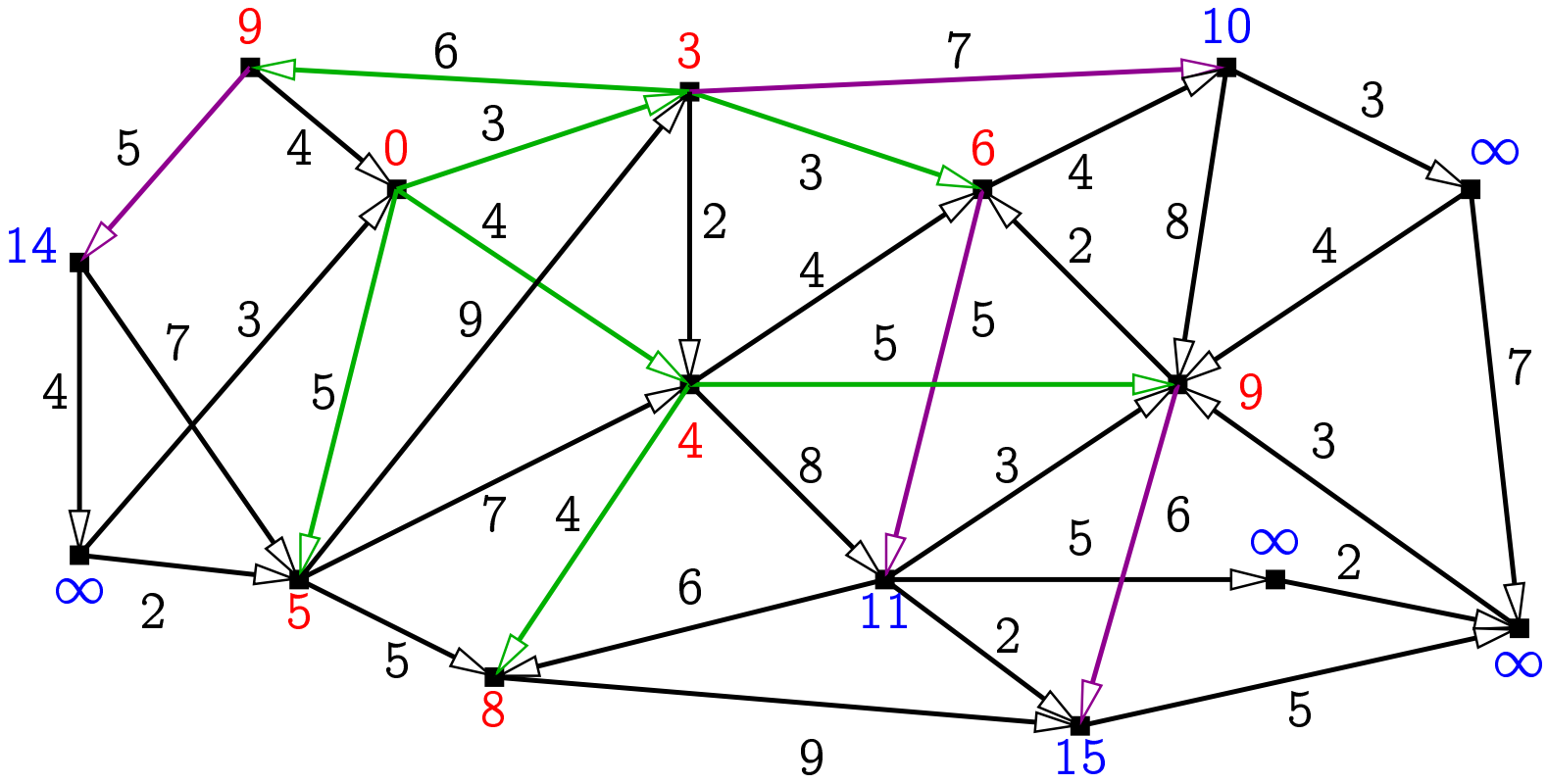
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



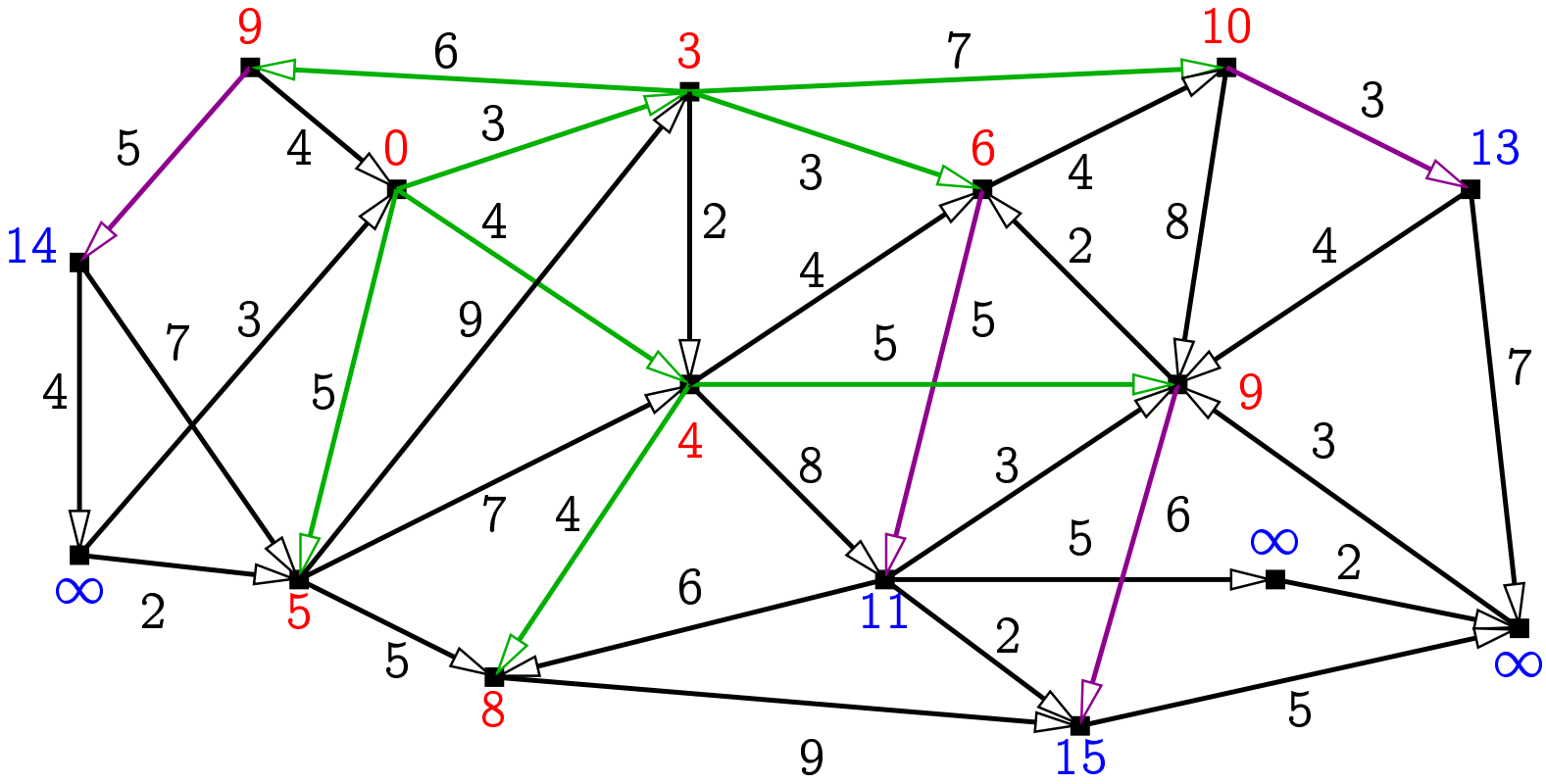
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



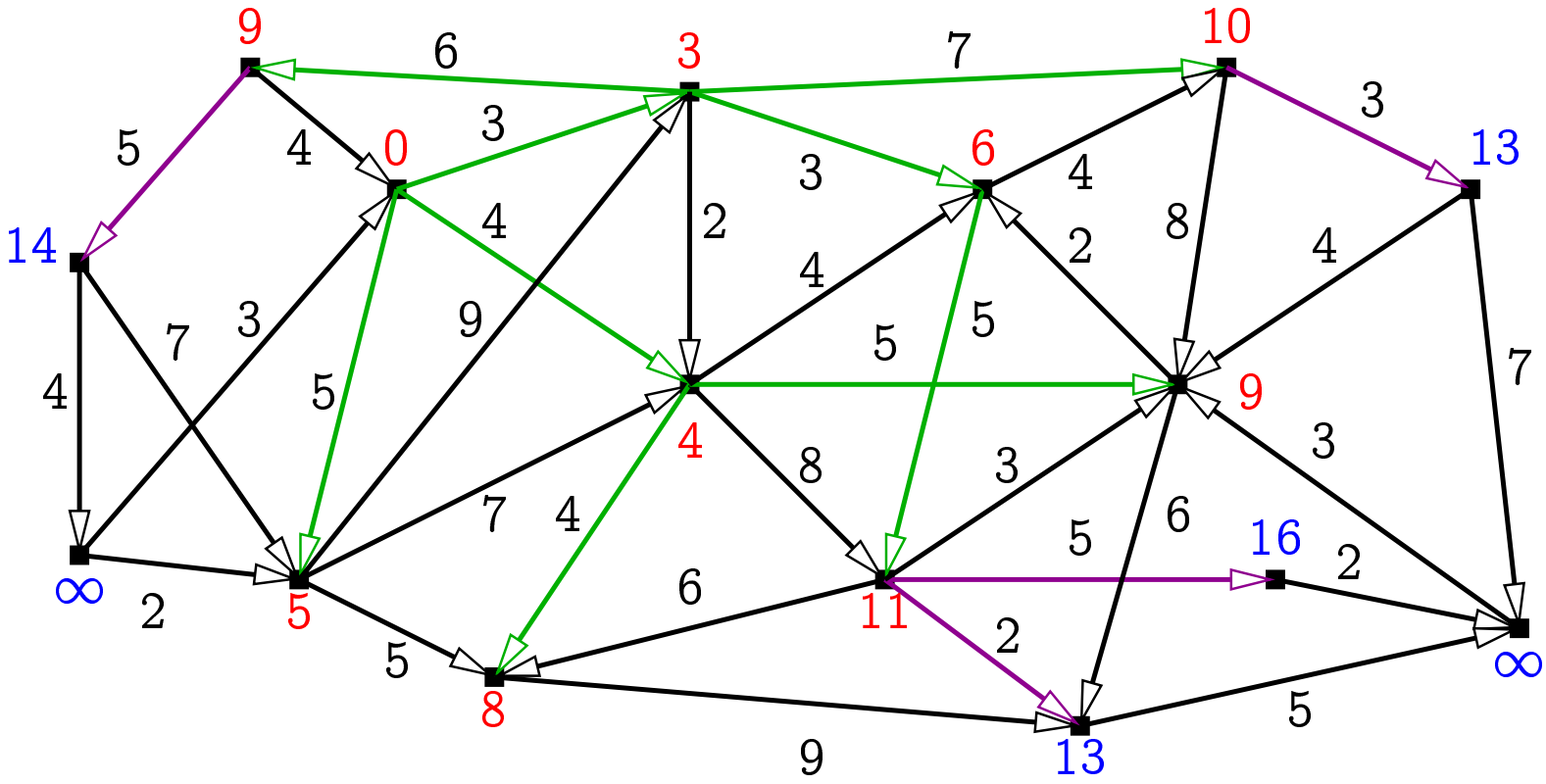
Näide: (V_1-e kuuluva tipu kaugus u-st, V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib, π-ga määratud servad)



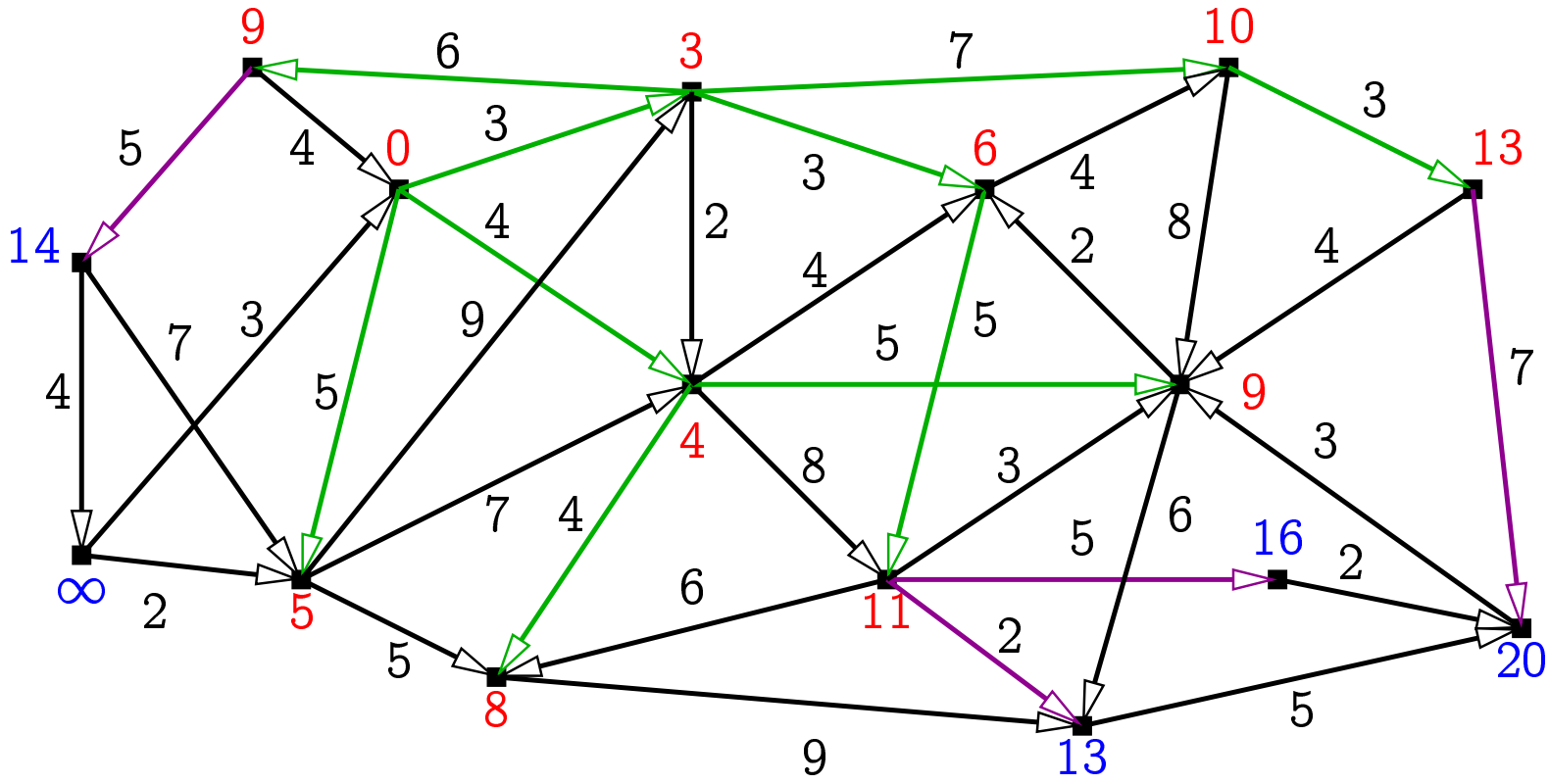
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



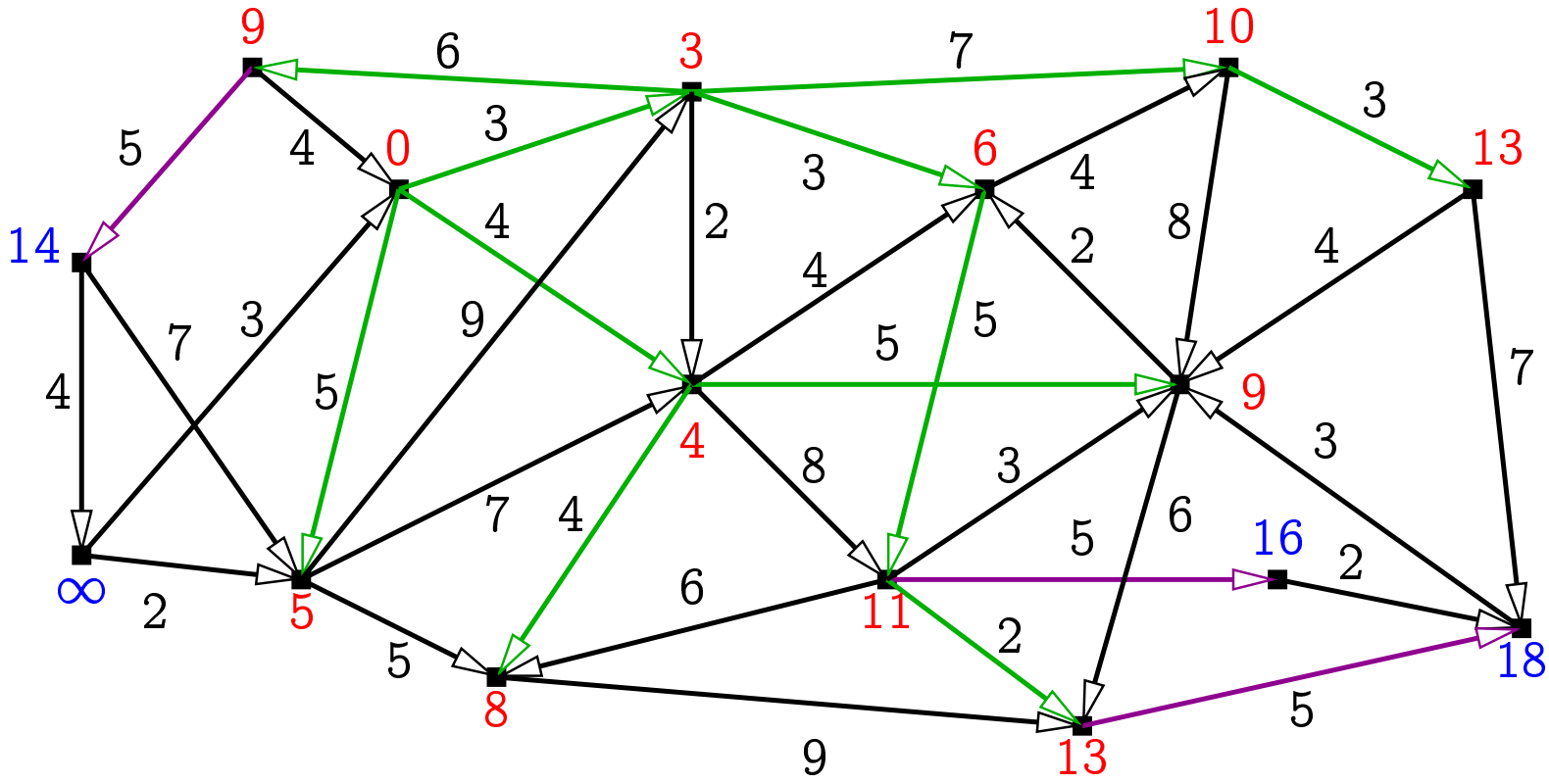
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



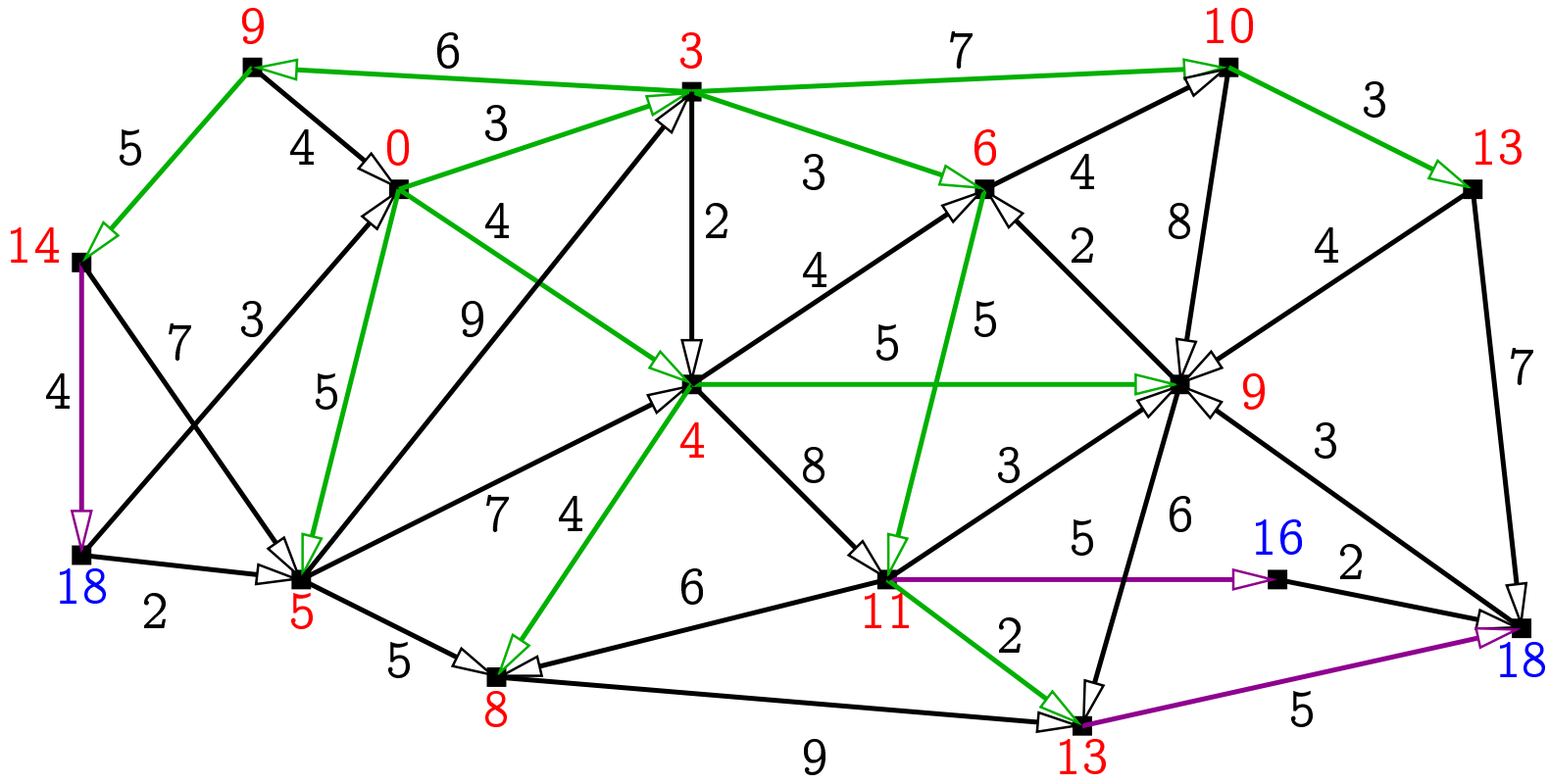
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



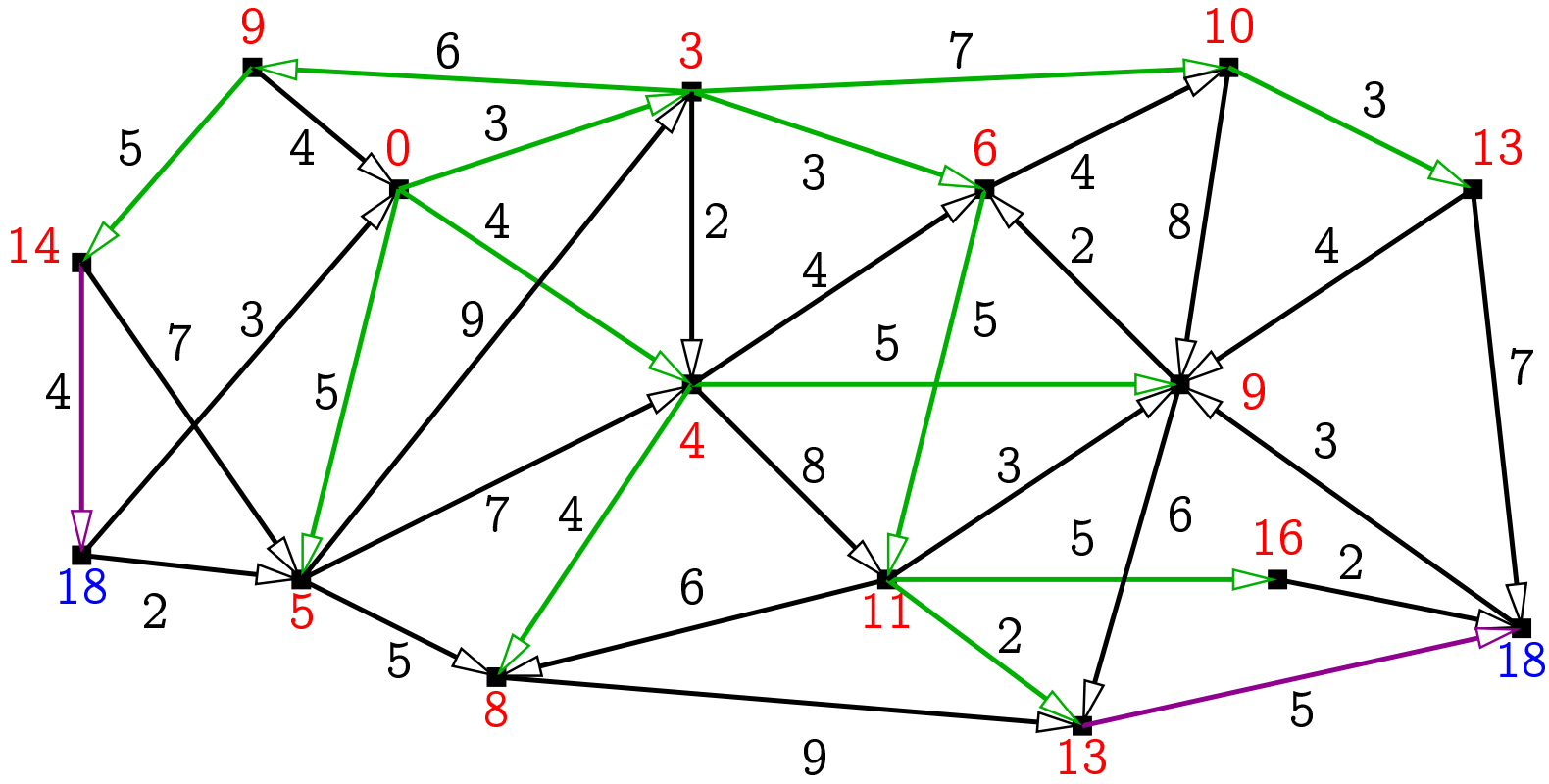
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



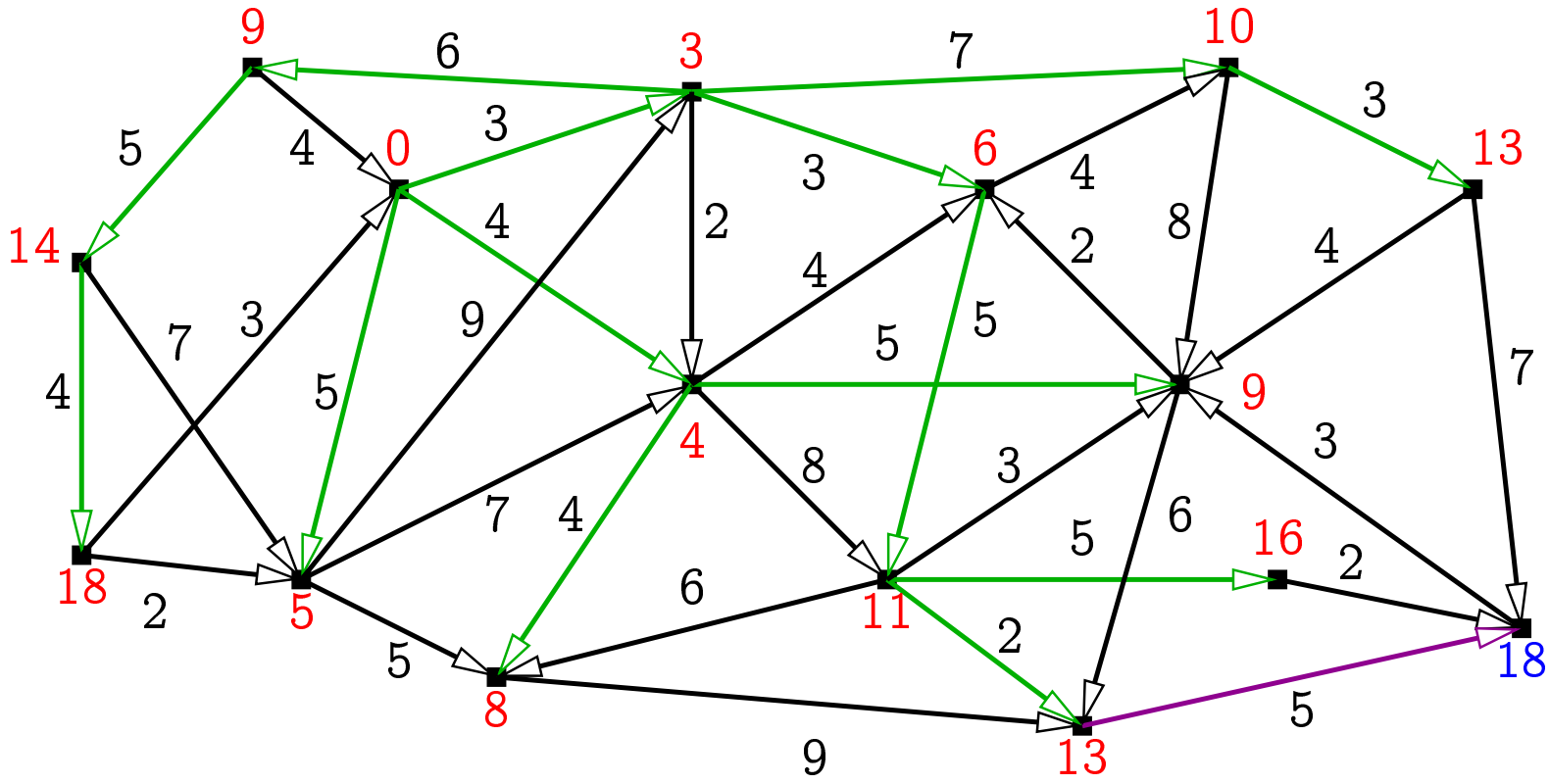
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



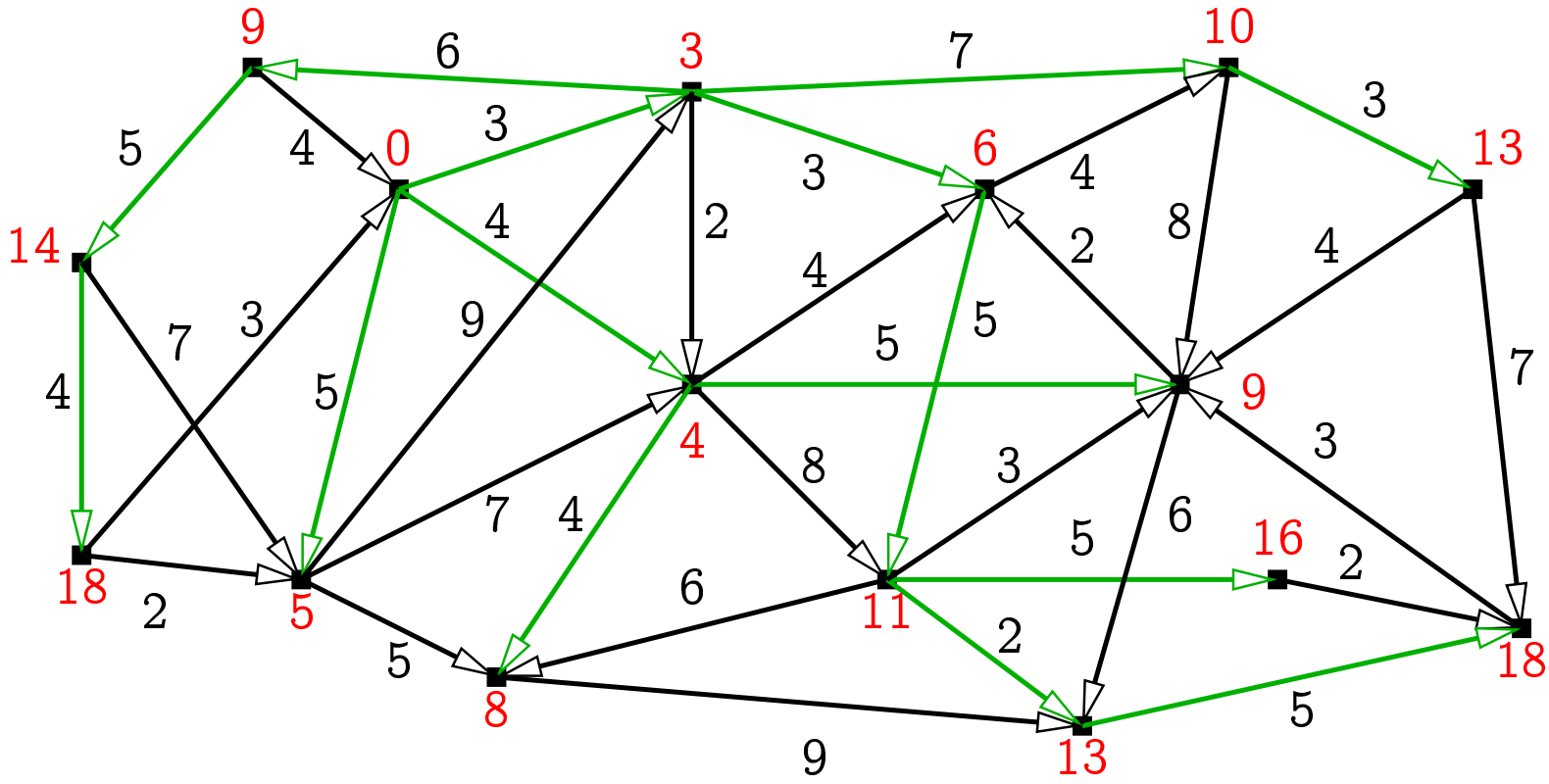
Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



Näide: (V_1-e kuuluva tipu kaugus u-st,V_2-e kuuluva tipu kaal ja serv, mis selle kaalu realiseerib,π-ga määratud servad)



Algoritmi realiseerides võib tipu kauguse u-st (V_1-e kuuluvate tippude jaoks) ja tipu kaalu (V_2-e kuuluvate tippude jaoks) kasutada sama massiivi D.

Kui me tõstame tipu w hulgast V_2 hulka V_1 ja seejuures tema naabertippude kaalusid vähendame, siis ei ole meil tarvis ilmutatult kontrollida, kas naabertipp kuulub V_2-e. Kui me saame naabertipu kaalu vähendada, siis oli tema kaal suurem kui w kaal ja seega kuulus ta V_2-e.

Olgu Q mingi kuhi (kahend- või Fibonacci). Lühimate teede leidmise algoritm on siis (nimetatakse *Dijkstra* algoritmiks):

```
 1  D[u] := 0
 2  for all v ∈ V\{u} do D[v] := ∞
 3  for all v ∈ Gu do D[v] := ℓ(u, v); π[v] := u
 4  Q := kuhjasta(V\{u})          --- D[·] järgi
 5  while Q ei ole tühi do
 6      w ⇐ Q
 7      if D[w] = ∞ then break
 8      for all v ∈ Gw do
 9          if D[v] > D[w] + ℓ(w, v) then
10              D[v] := D[w] + ℓ(w, v)
11              π[v] := w
12              vii_üles(Q, v)
13  return (D, π)
```

Keerukus:

Ridu 5–7 täidetakse $O(|V|)$ korda. Reas 6 on min. kaaluga elemendi võtmine kuhjast, keerukus $O(\log |V|)$, kõigi iteratsioonide peale kokku seega $O(|V| \log |V|)$.

Ridu 8–12 täidetakse $O(|E|)$ korda. Reas 12 on kuhjaparandusoperatsioon, keerukus $O(\log |V|)$ või $O(1)$, sõltuvalt kuhjast. Kõigi iteratsioonide peale kokku seega $O(|E| \log |V|)$ või $O(|E|)$.

Dijkstra algoritmi kogukeerukus on siis

- $O((|V| + |E|) \log |V|)$, kui kasutame kahendkuhja;
- $O(|V| \log |V| + |E|)$, kui kasutame Fibonacci kuhja.

Ülesanne: antud graaf $G = (V, E)$ koos servade pikkustega, nii et negatiivse pikkusega tsükleid ei ole. Leia kõikvõimalike $u, v \in V$ jaoks $d(u, v)$.

Kui negatiivse pikkusega servi ei ole, siis saame kasutada iga tipu jaoks Dijkstra algoritmi. Keerukus $O(|V|^2 \log |V| + |V||E|)$.

Kui on ka negatiivse pikkusega servi, siis Bellman-Fordi algoritmi kasutamine iga tipu jaoks oleks keerukusega $O(|V|^2|E|)$.

Järgnevas vaatame (üldiselt) efektiivsemaid algoritme kui Bellman-Ford iga tipu jaoks.

Tähistusi:

- $\mathrm{Mat}_{X\times Y}(K)$ — kõigi maatriksite hulk, mille read on indekseeritud hulga X elementidega, veerud hulga Y elementidega ja maatriksi elemendid on hulgast K.
- $\mathbb{R}_\infty$ — hulk $\mathbb{R}\cup\{\infty\}$.
- Kui $L\in\mathrm{Mat}_{X\times Y}(K)$, $i\in X$ ja $j\in Y$, siis $[L]_{ij}$ tähistagu maatriksi L elementi reas i ja veerus j.

Vaatame maatriksit $L \in \mathrm{Mat}_{V \times V}(\mathbb{R}_\infty)$, kus

$$[L]_{uv} = \begin{cases} 0, & \text{kui } u = v \\ \ell(u,v), & \text{kui } u \neq v \end{cases} .$$

Siis $[L]_{uv}$ näitab lühima sellise tee pikkust tipust u tippu v, kus on ülimalt üks serv.

Olgu $W_i \in \mathrm{Mat}_{V \times V}(\mathbb{R}_\infty)$ selline, et $[W_i]_{uv} = D_i(u,v)$. Siis $W_1 = L$ ning W_0 on selline, kus peadiagonaalil on 0-d ja mujal ∞-d.

Neg. pikkusega tsüklite puudumise tõttu $W_{|V|-1} = W_{|V|} = W_{|V|+1} = \cdots$. Meil tuleb see maatriks leida.

Vastavalt varemnäidatud seosele D_{i+1} ja D_i vahel:

$$[W_{i+1}]_{uv} = \min_{w \in V}([W_i]_{uw} + [L]_{wv}) \ .$$

kombineeri_maatriksid(A, B) on

```
1  for all u ∈ V do
2      for all v ∈ V do
3          C[u, v] := ∞
4          for all w ∈ V do
5              C[u, v] := min(C[u, v], A[u, w] + B[w, v])
6  return C
```

Siis $W_{i+1} =$ kombineeri_maatriksid(W_i, L).

kombineeri_maatriksid keerukus on $O(|V|^3)$.

initsialiseeri_L on

```
1  for all u ∈ V do
2      for all v ∈ V do
3          L[u, v] := ℓ(u, v)
4      L[u, u] := 0
5  return L
```

Siis tippude kauguste maatriksi võib leida järgmiselt:

```
1  L := initsialiseeri_L; W_1 := L
2  for i := 2 to |V| − 1 do
3      W_i := kombineeri_maatriksid(W_{i−1}, L)
4  return W_{|V|−1}
```

Keerukus $O(|V|^4)$, pole parem kui Bellman-Ford igale tipule.

Algebraline struktuur $(R, \overline{+}, \bar{\cdot}, \overline{0}, \overline{1})$ on *poolring*, kui

- $(R, \overline{+}, \overline{0})$ on kommutatiivne monoid;
- $(R, \bar{\cdot}, \overline{1})$ on monoid;
- $\overline{0}$ on korrutamise nullelement, s.t. $\overline{0}\bar{\cdot}r = r\bar{\cdot}\overline{0} = \overline{0}$ iga $r \in R$ jaoks;
- $\overline{+}$ distributeerub $\bar{\cdot}$ suhtes mõlemalt poolt. S.t. iga $a, b, c \in R$ jaoks

$$a\bar{\cdot}(b\overline{+}c) = a\bar{\cdot}b\overline{+}a\bar{\cdot}c$$
$$(a\overline{+}b)\bar{\cdot}c = a\bar{\cdot}c\overline{+}b\bar{\cdot}c \ .$$

Lause. Olgu R mingi poolring. Siis $\mathrm{Mat}_{V\times V}(R)$-s on maatriksite korrutamine assotsiatiivne.

Lause. $(\mathbb{R}_\infty, \min, +, \infty, 0)$ on poolring.

Selles poolringis $W_{i+1} = W_i \cdot L$. Seega $W_i = L^i$.

kombineeri_maatriksid kujutas endast just üle selle poolringi defineeritud maatriksite korrutamist.

Meie ülesandeks on leida L-i küllalt suur aste (vähemalt $|V|-1$). Assotsiatiivsuse tõttu võib seda leida L-i korduvalt ruutu tõstes.

```
1  L := initsialiseeri_L; W := L
2  for i := 1 to ⌈log |V|⌉ do
3      W := kombineeri_maatriksid(W, W)
4  return W
```

Keerukus: $O(|V|^3 \log |V|)$, mis on parem kui Bellman-Ford iga tipu jaoks, kui $|E| = \Omega(|V| \log |V|)$.

Oleme leidnud lühimate teede pikkused tippude vahel. Aga kuidas leida tegelikke teid?

Iga kahe tipu $u, v \in V$ jaoks oleks tarvis leida tipp w — eelviimane tipp lühimas tees u-st v-sse. S.t. otsime maatriksit $\Pi \in \mathrm{Mat}_{V \times V}(V)$.

Olgu tipp u fikseeritud. Olgu $D[v]$ lühima tee pikkus u-st v-sse. Näitame, kuidas leida massiiv π, nii et $\pi[v]$ oleks eelviimane tipp lühimas tees u-st v-sse.

Maatriksi Π leidmiseks kutsume siis seda protseduuri välja iga $u \in V$ jaoks.

Moodustame graafi G_u järgmiselt:

- G_u tippudeks on G tipud;
- G_u-s on serv w-st v-sse parajasti siis, kui G-s on serv w-st v-sse ja kui $D[v] - D[w] = \ell(w, v)$.

Graafis G_u on iga tee u-st v-sse mingi graafi G lühim tee u-st v-sse.

Graafi G_u laiuti / sügavuti / mingil muul viisil läbides leiame teed u-st teistesse tippudesse.

Olgu J mingi dünaamiline järjend. Järgmine protseduur leiab π.

```
1  for all v ∈ V do v.läbitud := false
2  J := ∅; J ⇐ u; u.läbitud := true
3  while J ei ole tühi do
4      w ⇐ J
5      for all v ∈ G_u w do
6          if ¬v.läbitud then
7              v.läbitud := true
8              π[v] := w
9              J ⇐ v
```

Keerukus:

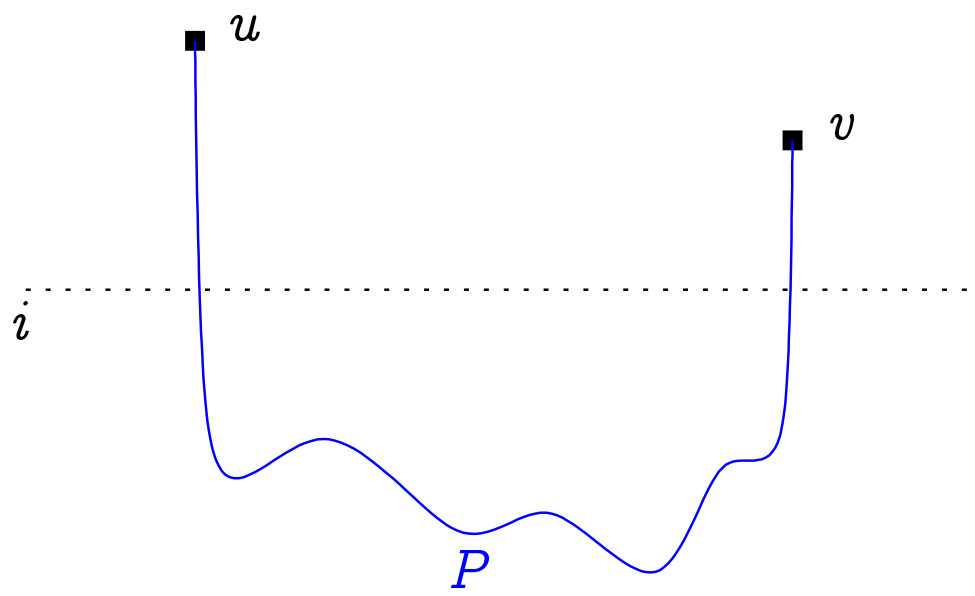
- G_u moodustamiseks: $O(|V| + |E|)$.
- π leidmiseks $O(|V|+|E|)$, kui dünaamilise järjendi operatsioonid on konstantse keerukusega.

Kokku seega $O(|V| + |E|)$. Kuna $|V|$ on $O(|E|)$, siis on π leidmise keerukus ka $O(|E|)$.

Kogu Π leidmine vajab seega aega $O(|V||E|)$.

Loeme nüüd, et graafi tippudeks on naturaalarvud $1, 2, \ldots, n$.

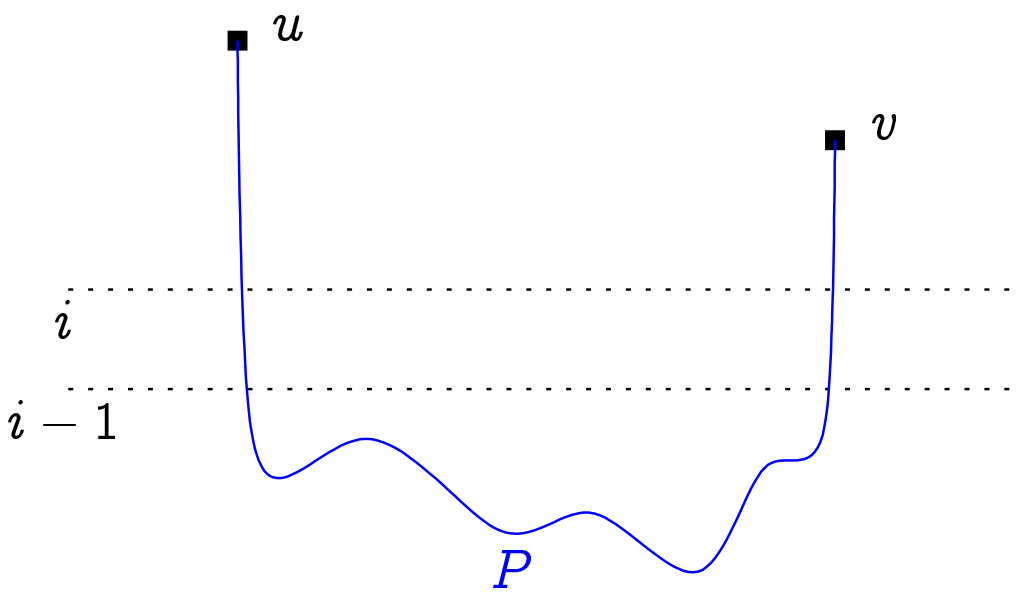
Iga $u, v, i \in V$ jaoks olgu $d^i(u, v)$ lühima sellise tee pikkus u-st v-sse, kus kõik tipud (peale u ja v) kuuluvad hulka $\{1, \ldots, i\}$. Siis $d(u, v) = d^n(u, v)$.



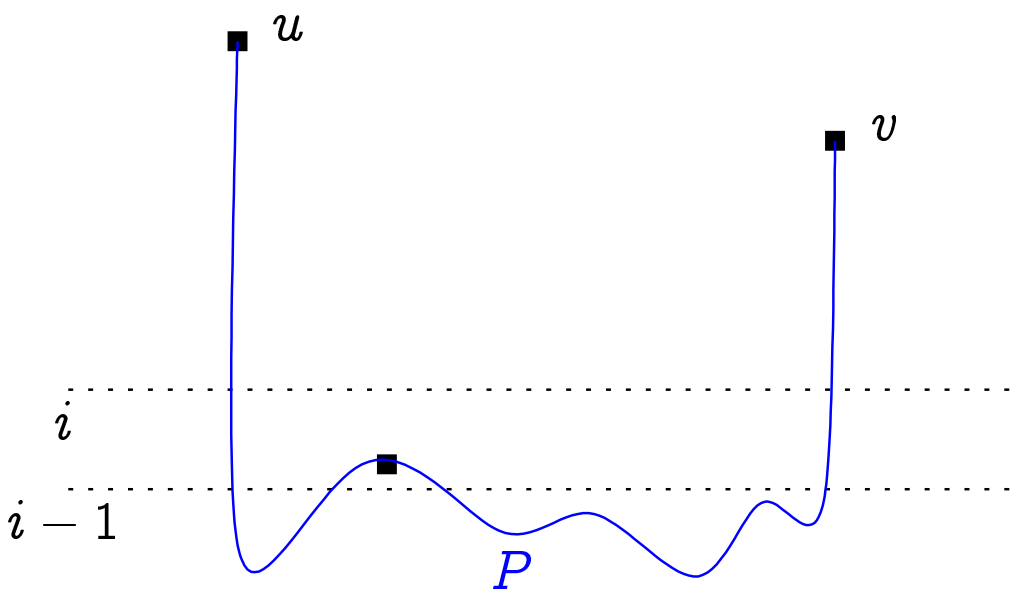
Defineerime veel $d^0(u, v) = \ell(u, v)$.

Millega võrdub $d^i(u,v)$. Tipp i võib antud tingimust rahuldavale lühimale teele kuuluda või mitte kuuluda.

Kui ei kuulu, siis $d^i(u,v) = d^{i-1}(u,v)$.



Kui kuulub, siis $d^i(u,v) = d^{i-1}(u,i) + d^{i-1}(i,v)$.



Tõepoolest, iga tipp kuulub lühimale ahelale ülimalt ühel korral. Seega on kõik teised tipud peale u, v ja i ahelal P hulgast $\{1, \ldots, i-1\}$.

Et liikuda võimalikult kiiresti tipust u tippu v, läbides seejuures ainult tippe $\{1, \ldots, i\}$ ja kindlasti läbides tippu i, tuleb

- liikuda võimalikult kiiresti tipust u tippu i, läbides seejuures ainult tippe $\{1, \ldots, i-1\}$,
- liikuda võimalikult kiiresti tipust i tippu v, läbides seejuures ainult tippe $\{1, \ldots, i-1\}$.

Kokkuvõttes:

$$d^i(u,v) = \min(d^{i-1}(u,v), d^{i-1}(u,i) + d^{i-1}(i,v)) \ .$$

Kõigi tippude vaheliste lühimate teede pikkuste leidmiseks leiame järjest $d^0, d^1, d^2, \ldots, d^n$.

Seda algoritmi nimetatakse *Floyd-Warshall*i algoritmiks.

```
1  for all u, v ∈ V do d[0][u, v] := ℓ(u, v)
2  for i := 1 to n do
3      for all u, v ∈ V do
4          d[i][u, v] := min(d[i − 1][u, v],
                             d[i − 1][u, i] + d[i − 1][i, v])
5  return d[n]
```

Keerukus — $O(|V|^3)$.

Me ei pea kogu massiivi d mälus hoidma. Igal hetkel läheb meil tarvis ainult $d[i]$-d ja $d[i-1]$-e. Seega piisab kahest $V \times V$ maatriksist.

Lühimad teed ise (s.t. — maatriksi Π) leiame nagu varem kirjeldatud, keerukusega $O(|V||E|)$.

Kui meil ei ole negatiivse pikkusega servi, siis saab kõigi tippude vaheliste kauguste leidmiseks kasutada Dijkstra algoritmi, rakendades seda iga tipu jaoks.

Tema keerukus $O(|V|^2 \log |V| + |V||E|)$ pole halvem kui Floyd-Warshalli algoritmi keerukus $O(|V|^3)$. Kui servi on vähe, s.t. $|E|$ on väiksem kui $\Theta(|V|^2)$, siis on Dijkstra algoritmi rakendamine iga tipu jaoks kiirem kui Floyd-Warshalli algoritm.

Järgnevas näitame, kuidas kasutada Dijkstra algoritmi ka juhul, kui graafis on negatiivse pikkusega servi. (*Johnson*i algoritm)

Tahame defineerida graafi servadel sellised uued pikkused $\hat{\ell}(v, w)$, et

- kui P on lühim tee u-st v-sse servade pikkuste ℓ järgi, siis on ta lühim tee u-st v-sse ka servade pikkuste $\hat{\ell}$ järgi;
- $\hat{\ell}(v, w) \geqslant 0$ iga $v, w \in V$ jaoks.

Kui oleme sellised pikkused defineerinud, siis kasutame Dijkstra algoritmi, kasutades pikkusi $\hat{\ell}$.

Olgu $h : V \longrightarrow \mathbb{R}$ suvaline funktsioon. Defineerime kõigi $u, v \in V$ jaoks $\hat{\ell}(u, v) = \ell(u, v) + h(u) - h(v)$.

Selliselt defineeritud $\hat{\ell}$ rahuldab esimest tingimust eelmiselt slaidilt. Tõepoolest, olgu $v_0 \to v_1 \to v_2 \cdots \to v_n$ mingi tee graafis G. Siis

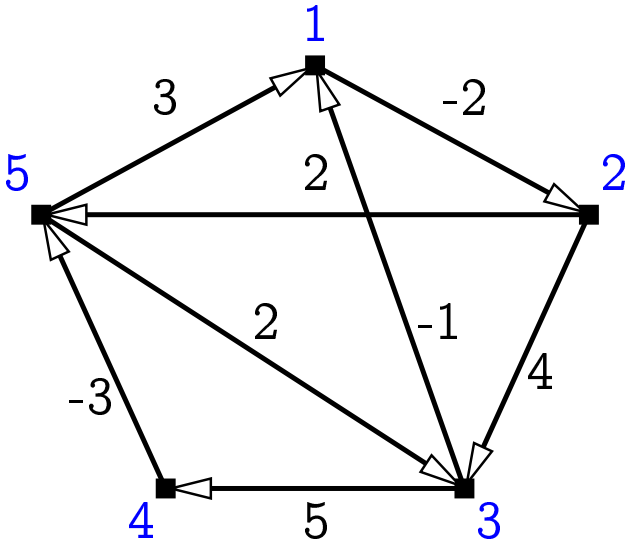
$$\hat{\ell}(P) := \sum_{i=1}^{n} \hat{\ell}(v_{i-1}, v_i) = \sum_{i=1}^{n} \big(\ell(v_{i-1}, v_i) + h(v_{i-1}) - h(v_i)\big) =$$

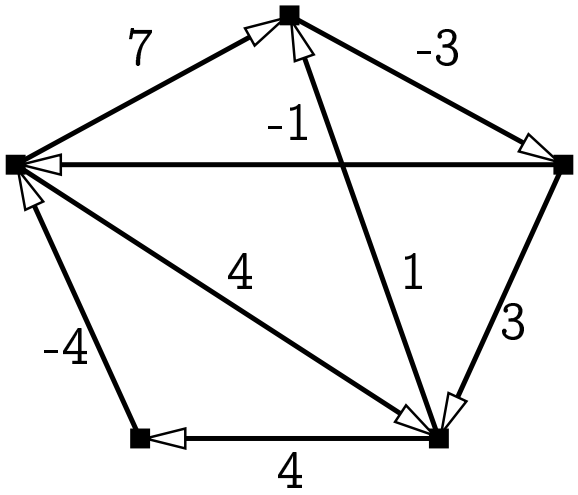
$$\sum_{i=1}^{n} \ell(v_{i-1}, v_i) + \sum_{i=0}^{n-1} h(v_i) - \sum_{i=1}^{n} h(v_i) =$$

$$\sum_{i=1}^{n} \ell(v_{i-1}, v_i) + h(v_0) - h(v_n) =: \ell(P) + h(v_0) - h(v_n) \ .$$

S.t. kõigi teede pikkus kahe fikseeritud tipu vahel muutub samapalju.

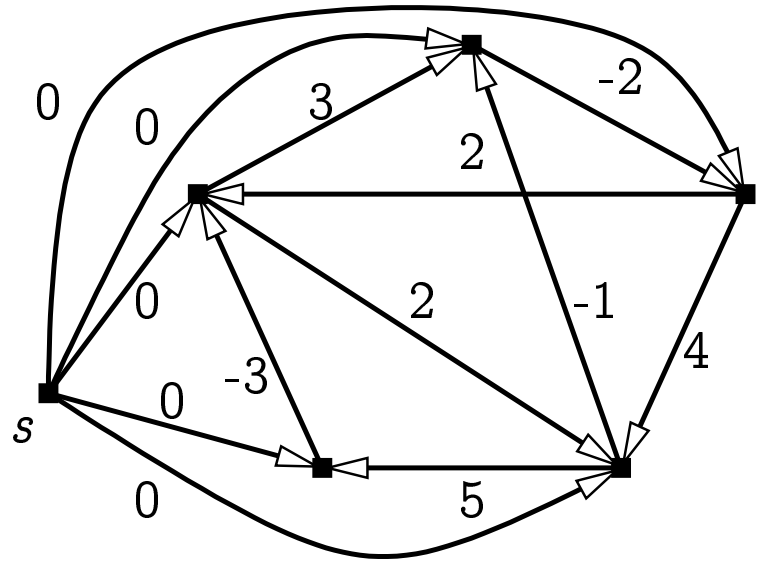
Näide ℓ-i muutmisest funktsiooni h abil.





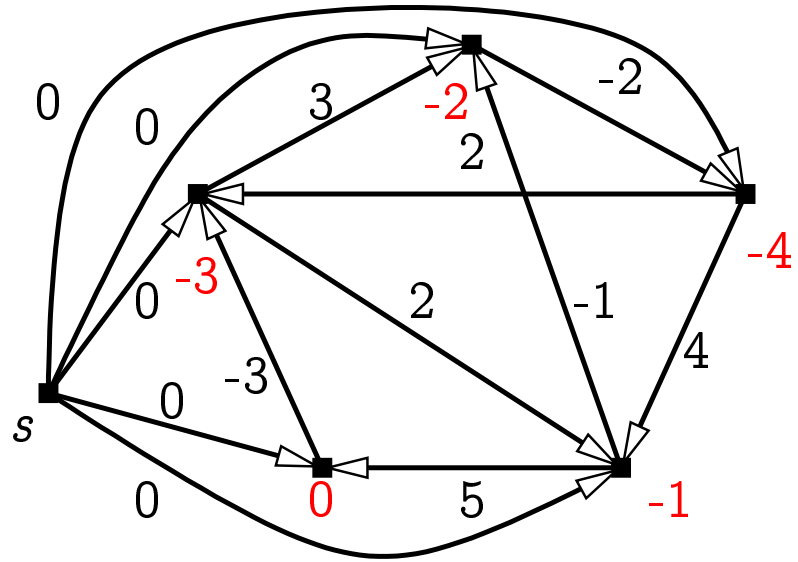
Me tahame h-i, mis kõik pikkused mittenegatiivseks muudaks.

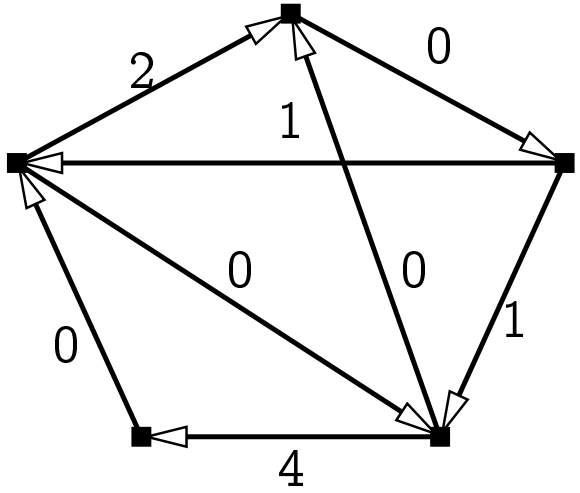
Selleks lisame graafile ühe täiendava tipu s ja servad tipust s kõigisse teistesse tippudesse. Nende servade pikkuseks võtame 0.



Seejuures ei lisandu graafi tsükleid, sest ühestki tipust ei saa minna s-i. Samuti jäävad kõigi olemasolevate tsüklite pikkused samaks.

Ka uues graafis pole negatiivse pikkusega tsükleid. Bellman-Fordi algoritmi abil leiame tipu s kauguse kõigist teistest tippudest.





Võtame $h(v) := d(s, v)$.

Olles defineerinud $h(v) = d(s, v)$, on meil

$$\hat{\ell}(u, v) = \ell(u, v) + d(s, u) - d(s, v) \ .$$

Meil on $d(s, v) \leqslant d(s, u) + \ell(u, v)$, seega $\hat{\ell}(u, v) \geqslant 0$.

Johnsoni algoritm:

- Lisa graafile täiendav tipp s, ühenda kõigi teiste tippudega 0-pikkusega servade abil.
- Leia Bellman-Fordi algortmiga s-i kaugus kõigist teistest tippudest.
- Esialgses graafis defineeri $\hat{\ell}(u,v) = \ell(u,v) + d(s,u) - d(s,v)$.
- Esialgse graafi iga tipu u jaoks leia Dijkstra algoritmiga tema kaugus ($\hat{\ell}$ järgi) kõigist teistest tippudest. Samuti leia lühimad teed ise.
- Leia lühimate teede pikkused ℓ järgi. Kui $\hat{\ell}$ järgi on u ja v kaugus $\hat{d}(u,v)$, siis $d(u,v) = \hat{d}(u,v) + d(s,v) - d(s,u)$.

Dijkstra algoritm töötleb tippe nende kauguse järjekorras u-st.

Kui meie eesmärgiks on tegelikult leida lühim tee tipust u mingisse konkreetsesse tippu t, siis võib algoritmis itereerimise lõpetada hetkel, kui kuhjast Q võetakse tipp t.

Kuhjast Q võetakse tippe nende kauguse järjekorras tipust u.

Et see hetk (t võtmine Q-st) saabuks kiiremini, võiksime defineerida servadel uued pikkused $\hat{\ell}(\cdot,\cdot)$, nii et $\hat{d}(u,t)$ oleks väike (ja lühimad teed jääksid samaks).

```
 1  D[u] := 0
 2  for all v ∈ V\{u} do D[v] := ∞
 3  for all v ∈ Gu do D[v] := ℓ(u,v); π[v] := u
 4  Q := kuhjasta(V\{u})          --- D[·] järgi
 5  while Q ei ole tühi do
 6      w ⇐ Q
 7      if D[w] = ∞∨w = t then break
 8      for all v ∈ Gw do
 9          if D[v] > D[w] + ℓ(w,v) then
10              D[v] := D[w] + ℓ(w,v)
11              π[v] := w
12              vii_üles(Q,v)
13  return (D,π)
```

Iga tipu v jaoks defineerime suuruse $h(v)$. Olgu $\hat{\ell}(v,w) = \ell(v,w) - h(v) + h(w)$.

Siis $\hat{d}(v,w) = d(v,w) - h(v) + h(w)$. Seega peaks $h(u)$ olema suur ja $h(t)$ olema väike. Loomulik on võtta $h(t) = 0$.

Et servade pikkused oleksid mittenegatiivsed, peab iga $v, w \in V$ jaoks kehtima

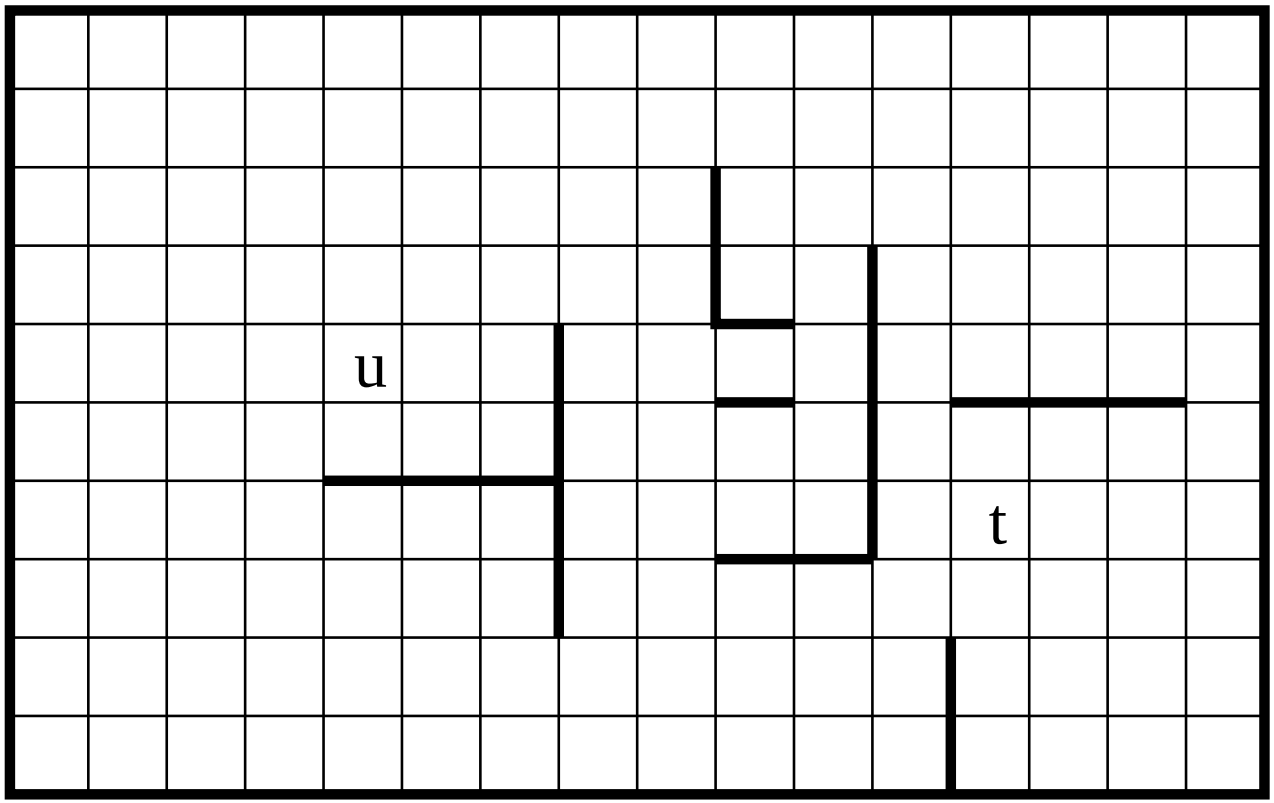
- $h(v) - h(w) \leqslant \ell(v,w)$ kui $w \in Gv$;
- $h(v) - h(w) \leqslant d(v,w)$.
  - Kui $h(t) = 0$, siis $h(v) \leqslant d(v,t)$.

Kui graafiks on mingid punktid ja ühendused maastikul, siis $h(v)$-ks sobib v kaugus t-st linnulennul.

```
 1  D[u] := h(u)
 2  for all v ∈ V\{u} do D[v] := ∞
 3  for all v ∈ Gu do D[v] := ℓ(u,v)+h(v); π[v] := u
 4  Q := kuhjasta(V\{u})          --- D[·] järgi
 5  while Q ei ole tühi do
 6      w ⇐ Q
 7      if D[w] = ∞∨w = t then break
 8      for all v ∈ Gw do
 9          if D[v]−h(v) > D[w]−h(w) + ℓ(w,v) then
10              D[v] := D[w]−h(w) + ℓ(w,v)+h(v)
11              π[v] := w
12              vii_üles(Q,v)
13  return (D,π)          --- meid huvitab ainult tee t-sse
```

See on A* lühima tee otsimise algoritm.

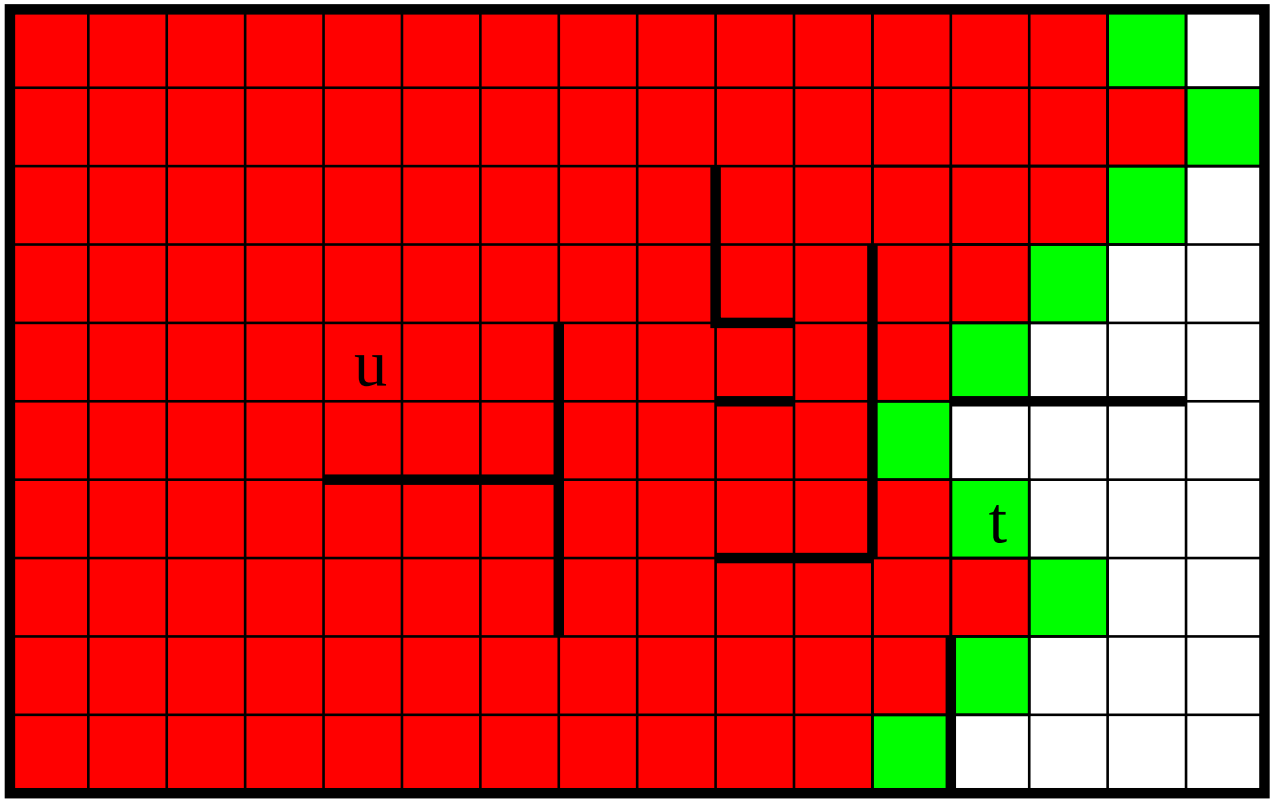
Näide. Vaatame graafi:



Tipud — ruudud. Iga ruut seotud nelja naabriga (v.a. läbi paksude joonte). Kõigi servade pikkus sama.

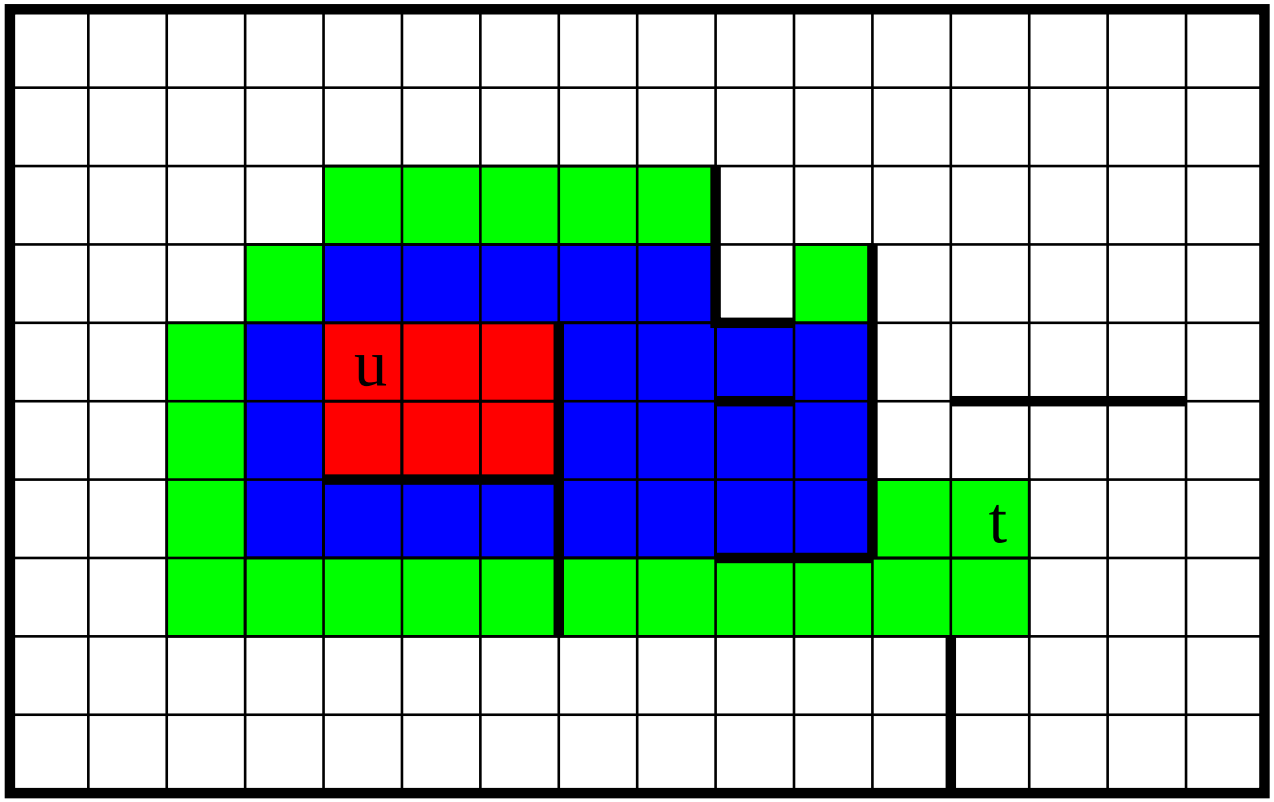
Olgu $h(v)$ võrdne $d(v,t)$-ga, kui sisemisi seinu ei oleks.

Dijkstra algoritm vaatab läbi järgmised ruudud:



(u-le lähemal ja sama kaugel kui t)

A* vaatab läbi järgmised ruudud:



$(d(u,v) + h(v) = 10 \quad 12 \quad 14)$