Paljudes rakendustes on tarvis dünaamilist kirjete hulka, mis toetaks kõigest järgmisi kolme operatsiooni:

- lisa kirje
- otsi kirjet (võtme järgi)
- kustuta kirje (võtme järgi)

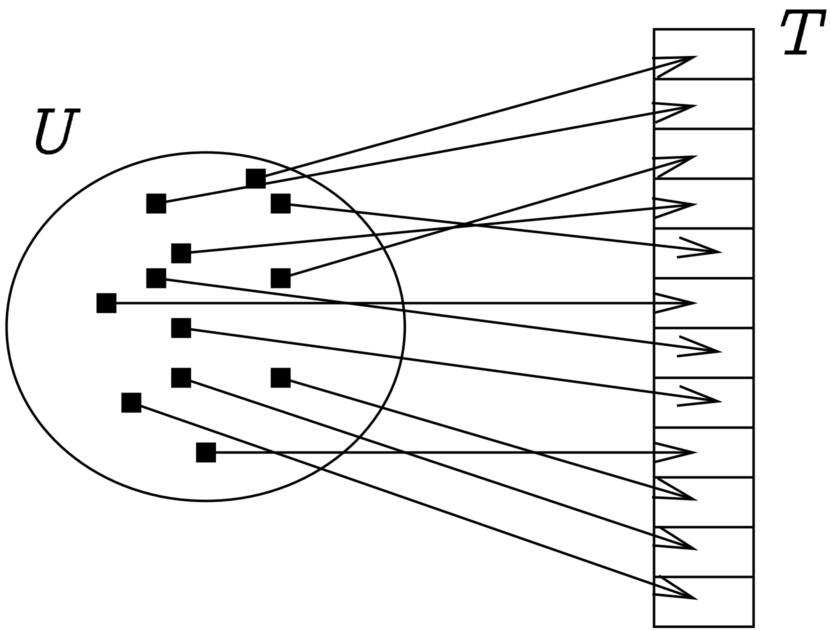
Need operatsioonid võiksid võtta nii vähe aega kui võimalik.

Kui võimalike võtmeväärtuste arv on väike, siis sobib selleks andmestruktuuriks massiiv.

Olgu võimalike võtmete hulk $U = \{0, 1, \ldots, m-1\}$. Siis võtame andmestruktuuriks massiivi pikkusega m.

Massiivi elementidel peab olema lisaväli $.t\ddot{u}hi$. Alguses on see kõigil elementidel tõene. Väli näitab, kas vastava võtmega kirje on olemas.

Operatsioonide keerukus: $O(1)$.


U
T

Kui $|U|$ on suur, võtab massiiv liiga palju ruumi.

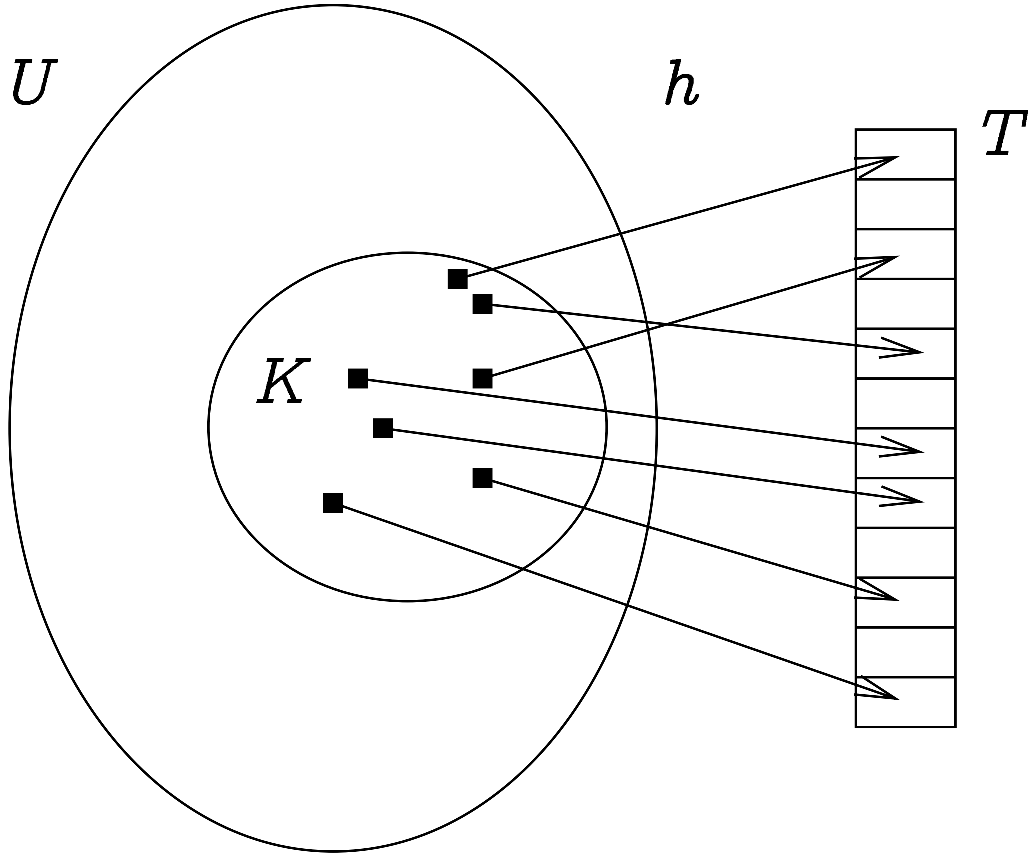
Kui tegelikke kirjeid on alati palju vähem kui võimalike võtmeid, siis raiskab massiiv liiga palju mälu.

Võtame massiivi $T[0..m-1]$, nii et tegelike kirjete arv pole kunagi suurem kui m.

Defineerime mingil viisil *räsifunktsiooni* $h : U \to \{0, \ldots, m-1\}$.

Kirje võtmega k salvestame massiivielementi $T[h(k)]$. Võtit k otsime massiivielemendist $T[h(k)]$.

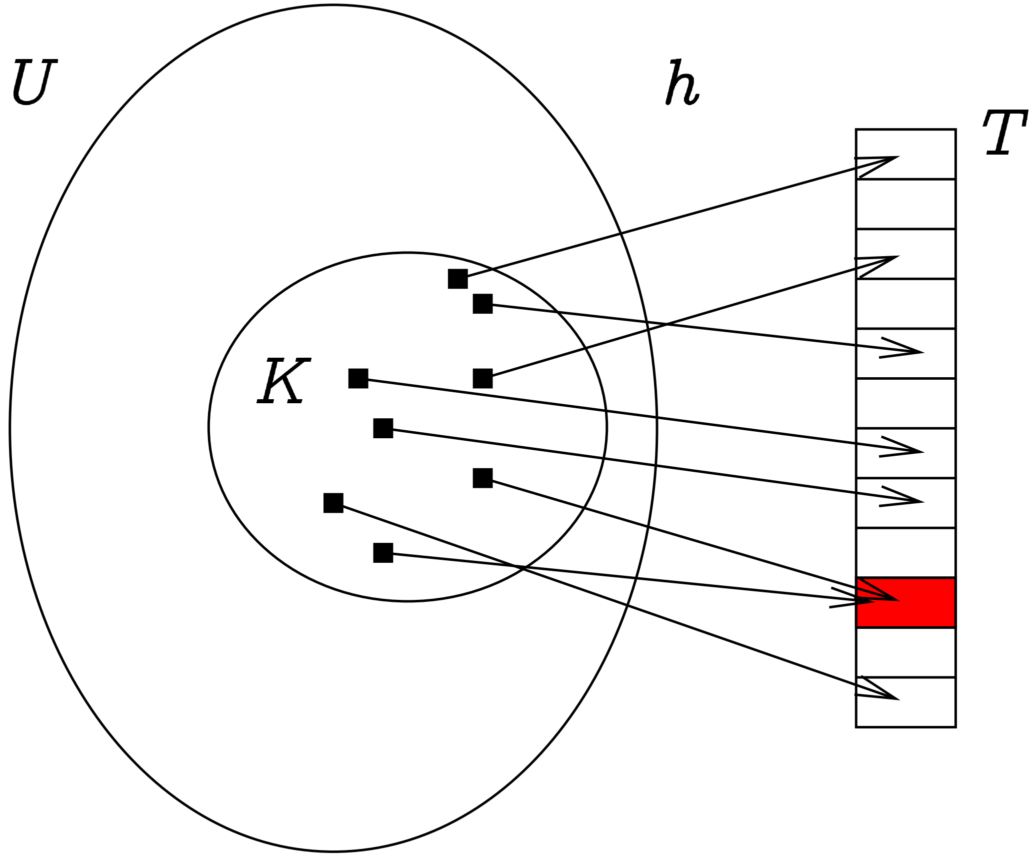
Sel viisil realiseeritud andmestruktuuri nimetatakse *paisktabeliks*.



U
h
T
K

Mis saab siis, kui tahame tabelisse lisada kahte (erineva võtmega) kirjet, mille võtmete räsid on võrdsed?

Seda olukorda nimetame *kollisiooniks* ehk *põrkeks*.



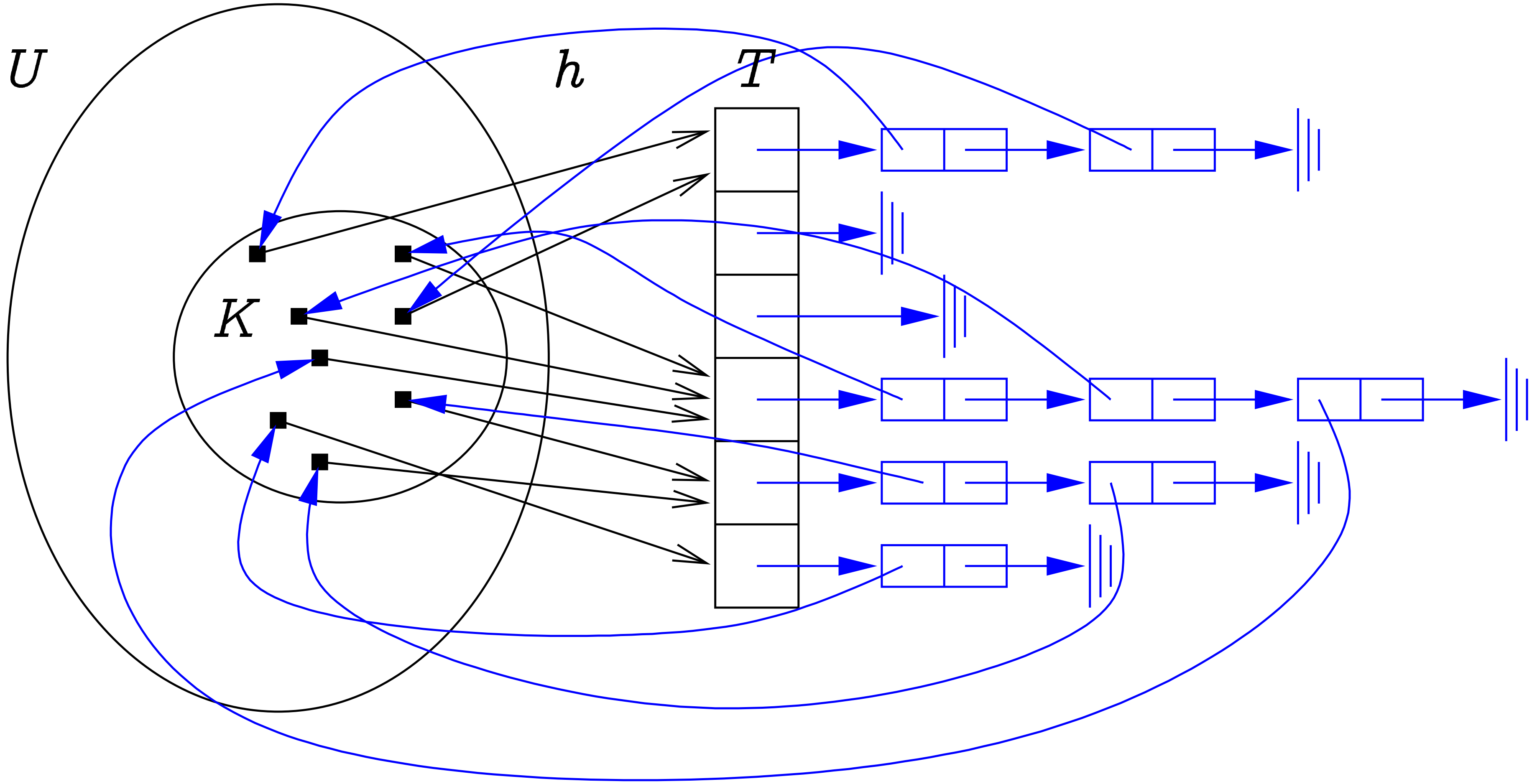
U
h
T
K

Kaks meetodit kollisioonide lahendamiseks:

- Välisahelate meetod: $T[l]$ viitab lihtahelale, mis sisaldab kõiki kirjeid, mille võtmete räsiks on l.

- Lahtise adresseerimise meetod: Võtmega k kirje lisamisel, kui $T[h(k)]$ on juba hõivatud, siis otsitakse sellele kirjele mingi teine koht massiivis T.

Välisahelate meetod



- Kirje võtmega k lisamine: lisame kirje ahelasse $T[h(k)]$.
  - Keerukus: $O(1)$.
- Võtme k otsimine: otsime lihtahelast, millele viitab $T[h(k)]$.
  - Keerukus: proportsionaalne selle ahela pikkusega.
- Kirje võtmega k kustutamine: kustutame lihtahelast, millele viitab $T[h(k)]$.
  - Keerukus: proportsionaalne selle ahela pikkusega.

Kui pikk on lihtahel, millele viitab $T[h(k)]$?

Halvimal juhul on tegelikult esinevate võtmete hulk selline, et nende kõigi räsid on samad.

Järgnevalt analüüsime otsimise ja kustutamise keerukust eeldusel, et juhuslikult valitud tegelikult esineva võtme räsi on ühtlaselt ning teiste võtmete räsidest sõltumatult jaotatud (hulgal $\{0, \ldots, m-1\}$).

Olgu α tabeli T *täituvus*, s.t. $\alpha = n/m$, kus n on tabelis olevate kirjete arv.

Tehtud eeldusel („lihtne ühtlane räsimine“) on ühe ahela keskmine pikkus α.

Lause. Lihtsa ühtlase räsimise korral võtab ebaõnnestunud otsimine aega $\Theta(1 + \alpha)$.

Tõestus. Tabelis mitteesineva võtme k räsi arvutamine võtab aega ühe ajaühiku ja tema otsimine ahelast, millele viitab $T[h(k)]$, $n_{h(k)}$ ajaühikut, kus n_i on selle lihtahela pikkus, millele viitab $T[i]$.

$n_{h(k)}$ keskmine väärtus on α. $\square$

Lause. Lihtsa ühtlase räsimise korral võtab tabelis oleva juhuslikult valitud võtme otsimine aega $\Theta(1+\alpha)$.

Tõestus. Olgu otsitav võti k lisatud tabelisse järjekorras i-ndana (iga konkreetse i tõenäosus on seejuures $1/n$).

Uuemad elemendid lisatakse lihtahelate algustesse, seega tuleb kirje võtmega k leidmiseks läbi vaadata need hiljemlisatud kirjet, mille võtmete räsi on $h(k)$.

Hiljemlisatud kirjeid on $n-i$. Lihtsa ühtlase räsimise tõttu on neist keskmiselt $\frac{n-i}{m}$ sellist, mille võtme räsi on $h(k)$. Nende keskmine (üle i) on

$$\sum_{i=1}^{n} \frac{1}{n} \cdot \frac{n-i}{m} = \frac{1}{nm} \left(\sum_{i=1}^{n} n - \sum_{i=1}^{n} i \right) = \frac{n}{m} - \frac{n+1}{2m} = \frac{n}{2m} - \frac{1}{2m} = \frac{\alpha}{2} - \frac{\alpha}{2n} \ .$$

Lisaks neile elementidele tuleb veel läbi vaadata kirje võtmega k ise. Enne otsimist tuleb arvutada $h(k)$. $\square$

Näiteid räsifunktsioonidest.

Hea räsifunktsioon peaks tagama, et lihtsa ühtlase räsimise eeldus on enam-vähem tagatud.

Kui võtmed on ühtlaselt ja sõltumatult jaotunud reaalarvud poollõigust $[0, 1)$, siis $h(k) := \lfloor km \rfloor$ tagab lihtsa ühtlase räsimise.

Edaspidi eeldame, et võtmed on täisarvud.

Näide: $h(k) := k \mod m$.

Et $h(k)$ sõltuks k kõikidest bittidest, siis ei tohiks m olla lähedal mõnele kahe astmele. Ka on hea, kui m on algarv.

Näide: Olgu $0 < A < 1$ fikseeritud. $h(k) := \lfloor m(kA - \lfloor kA \rfloor) \rfloor$.

Siin m-i väärtus pole nii kriitiline. A väärtus mõjutab h headust. Knuth soovitab võtta $A = (\sqrt{5} - 1)/2$.

Senini eeldasime, et paisktabelisse lisatavad võtmed on hästi jaotatud.

Kui me seda eeldust teha ei taha, aga paisktabeli operatsioonide keskmist keerukust ikkagi uurida tahame, siis tuleb meil juhuslikkus kuskile mujale sisse tuua.

Valime räsifunktsiooni juhuslikult (mingist räsifunktsioonide perest).

Järgnevas uurime, milline peaks olema sobiv pere.

Olgu $\mathcal{H}$ mingi lõplik funktsioonide hulk, mille määramispiirkonnaks on U ja muutumispiirkonnaks $\{0,\ldots,m-1\}$.

$\mathcal{H}$ on *universaalne* räsifunktsioonide hulk, kui

- iga $k, l \in U$ jaoks, kus $k \neq l$,
- ühtlaselt juhuslikult valitud $h \in \mathcal{H}$ jaoks
- $\Pr[h(k) = h(l)] \leqslant 1/m$.

Teisisõnu, iga $k, l \in U$, $k \neq l$ jaoks leidub ülimalt $|\mathcal{H}|/m$ sellist funktsiooni $h \in \mathcal{H}$, mille korral $h(k) = h(l)$.

Teoreem. Kui räsifunktsioon $h \in \mathcal{H}$ on ühtlaselt juhuslikult valitud universaalsest räsifunktsioonide perest, siis lihtahela, millele viitab $T[h(k)]$, keskmine pikkus on

- $\leqslant \alpha$, kui võti k pole tabelis;
- $\leqslant 1 + \alpha$, kui võti k on tabelis.

Tõestus. Iga tabelis oleva võtme $l \neq k$ on tõenäosus, et $h(k) = h(l)$, ülimalt $1/m$. Seega on $T[h(k)]$ keskmine pikkus, arvestamata võtit k, ülimalt $(n-1)/m$.

Kui k pole tabelis, siis on $T[h(k)]$ keskmine pikkus $(n - 1)/m < \alpha$.
Kui k on tabelis, siis on $T[h(k)]$ keskmine pikkus $(n - 1)/m + 1 < 1 + \alpha$. $\square$

Kuidas konstrueerida universaalset räsifunktsioonide hulka?

Olgu p mingi algarv, nii et $U \subseteq \mathbb{Z}_p$, kus $\mathbb{Z}_p = \{0, \ldots, p-1\}$. Olgu $\mathbb{Z}_p^* = \{1, \ldots, p-1\}$.

Fakt. Fikseerime $a \in \mathbb{Z}_p^*$. Siis kujutus $x \mapsto ax \mod p$ on hulga $\mathbb{Z}_p^*$ bijektsioon.

Fakt. Iga $x \in \mathbb{Z}_p^*$ jaoks leidub selline $y \in \mathbb{Z}_p^*$, et $xy \equiv 1 \pmod{p}$. Seda y-t tähistame x^{-1}.

Olgu $a \in \mathbb{Z}_p^*$ ja $b \in \mathbb{Z}_p$. Olgu $h_{a,b}(k) = ((ak + b) \mod p) \mod m$. Vaatame peret

$$\mathcal{H}_{p,m} = \{h_{a,b} : a \in \mathbb{Z}_p^*, b \in \mathbb{Z}_p\}$$

Teoreem. $\mathcal{H}_{p,m}$ on universaalne räsifunktsioonide hulk.

Tõestus. Olgu $k, l \in \mathbb{Z}_p$, kus $k \neq l$, ning $h_{a,b} \in \mathcal{H}_{p,m}$. Vaatame suurusi

$$r = ak + b \pmod{p}$$
$$s = al + b \pmod{p} \ .$$

Siis $r \neq s$, sest $ak \neq al \pmod{p}$.

Kui $r \neq s$, siis leiduvad a ja b, mis fikseeritud k ja l ($k \neq l$) jaoks selle r-i ja s-i annavad. Need on

$$a = (r - s) \cdot (k - l)^{-1} \pmod{p}$$
$$b = r - ak \pmod{p} \ .$$

Tähistame: $\mathbb{Z}_p^{2\neq} = \mathbb{Z}_p \times \mathbb{Z}_p \backslash \{(x,x) \,:\, x \in \mathbb{Z}_p\}$.

Iga $(k,l) \in \mathbb{Z}_p^{2\neq}$ defineerib bijektsiooni $\mathbb{Z}_p^* \times \mathbb{Z}_p$ ja $\mathbb{Z}_p^{2\neq}$ vahel. See on $(a,b) \mapsto (ak+b, al+b) \pmod p$.

Tõenäosus, et $h_{a,b}(k) = h_{a,b}(l)$, kui $h_{a,b} \in \mathcal{H}_{p,m}$ on valitud ühtlaselt juhuslikult, on tõenäosus, et $r \mod m = s \mod m$, kui $(r,s) \in \mathbb{Z}_p^{2\neq}$ on valitud ühtlaselt juhuslikult.

Fikseerime r-i. Tõenäosus, et juhuslikult valitud $s \in \mathbb{Z}_p \backslash r$ korral $r \mod m = s \mod m$, on ülimalt

$$\frac{\lceil p/m \rceil - 1}{p-1} \leqslant \frac{\frac{p+m-1}{m} - 1}{p-1} \leqslant \frac{\frac{p+m-1-m}{m}}{p-1} = \frac{1}{m} \quad .$$

□

Lahtise adresseerimise meetod.

Kõiki kirjeid hoitakse massiivis endas. Väli $.t\ddot{u}hi$ näitab, kas mingis massiivielemendis on kirje või ei. $\alpha \leqslant 1$.

Kui mingi kirje lisamisel on teine kirje juba ees, siis leiame järgmise koha, kuhu see kirje lisada, kuni lõpuks koha leiame.

Räsifunktsioon: $h : U \times \{0, \ldots, m-1\} \to \{0, \ldots, m-1\}$. Siin $h(k, i)$ on elemendi indeks, kuhu üritame lisada kirjet võtmega k, kui elemendid $T[h(k, 0)], \ldots, T[h(k, i-1)]$ on kõik juba hõivatud.

Iga $k \in U$ jaoks peab $h(k, \cdot)$ olema permutatsioon.

Kirje võtmega k lisamine:

- Arvuta $h(k,0)$, $h(k,1)$, $h(k,2)$,..., kuni on leitud selline i, et $T[h(k,i)].t\ddot{u}hi$ on tõene.
- Lisa kirje pessa $T[h(k,i)]$ ja muuda väli $.t\ddot{u}hi$ vääraks.

Võtme k otsimine:

- Arvuta $h(k,0)$, $h(k,1)$, $h(k,2)$,..., kuni on leitud selline i, et üks järgmistest tingimustest on tõene:
  1. $T[h(k,i)].v\tilde{o}ti = k$;
  2. $T[h(k,i)].t\ddot{u}hi$ on tõene.
- Tagasta leitud kirje (1. juhul) või NIL (2. juhul).

Kustutamine: tee uus tabel ja kopeeri sinna kõik kirjed, v.a. kustutatav.

Viise $h(k, i)$ defineerimiseks:

- $h(k, i) = (h'(k) + i) \mod m$;
  - kui $h(k, i) = h(l, j)$, siis ka $h(k, i + \delta) = h(l, j + \delta)$
- $h(k, i) = (h'(k) + c_1 i + c_2 i^2) \mod m$;
  - kui $h(k, 0) = h(l, 0)$, siis $h(k, \delta) = h(l, \delta)$
  - c_1, c_2, m peavad olema õigesti valitud
- topeltpaiskamine: $h(k, i) = (h_1(k) + ih_2(k)) \mod m$.
  - $h_2(k)$ ja m peavad olema ühistegurita
    * tavaliselt võetakse m-ks algarv

Järgnevas analüüsis teeme *ühtlase räsimise* eelduse:

> Ühtlaselt juhuslikult valitud $k \in U$ korral on $h(k, \cdot)$ ühtlaselt juhuslikult valitud permutatsioon m-elemendilisel hulgal.

Ükski eelmisel kilel toodud räsifunktsioonidest seda eeldust ei rahulda. Topeltpaiskamine on kõige lähemal.

Teoreem. Ühtlase räsimise eeldusel ning eeldusel, et paisktabelisse lisatud võtmed on valitud ühtlaselt ja üksteisest sõltumatult on massiivi T keskmine uuritud elementide arv ebaõnnestunud otsimisel ülimalt $1/(1-\alpha)$.

Tõestus. Tõenäosus, et otsimisel uuritakse vähemalt j-i massiivielementi, on

$$P_j = \frac{n}{m} \cdot \frac{n-1}{m-1} \cdots \frac{n-j+2}{m-j+2} \leqslant \left(\frac{n}{m}\right)^{j-1} = \alpha^{j-1}$$

Uuritud elementide arv on seega

$$\sum_{j=1}^{\infty} j(P_j - P_{j+1}) = \sum_{j=1}^{\infty} P_j \leqslant \sum_{i=0}^{\infty} \alpha^i = \frac{1}{1-\alpha} \quad .$$

$\square$

Kirje lisamine tabelisse vajab sama paljude elementide läbivaatust, kui ebaõnnestunud otsimine.

Teoreem. Ühtlase räsimise eelduse ning eeldusel, et paisktabelisse lisatud võtmed on valitud ühtlaselt ja üksteisest sõltumatult on massiivi T keskmine uuritud elementide arv õnnestunud otsimisel ülimalt $\frac{1}{\alpha} \ln \frac{1}{1-\alpha}$.

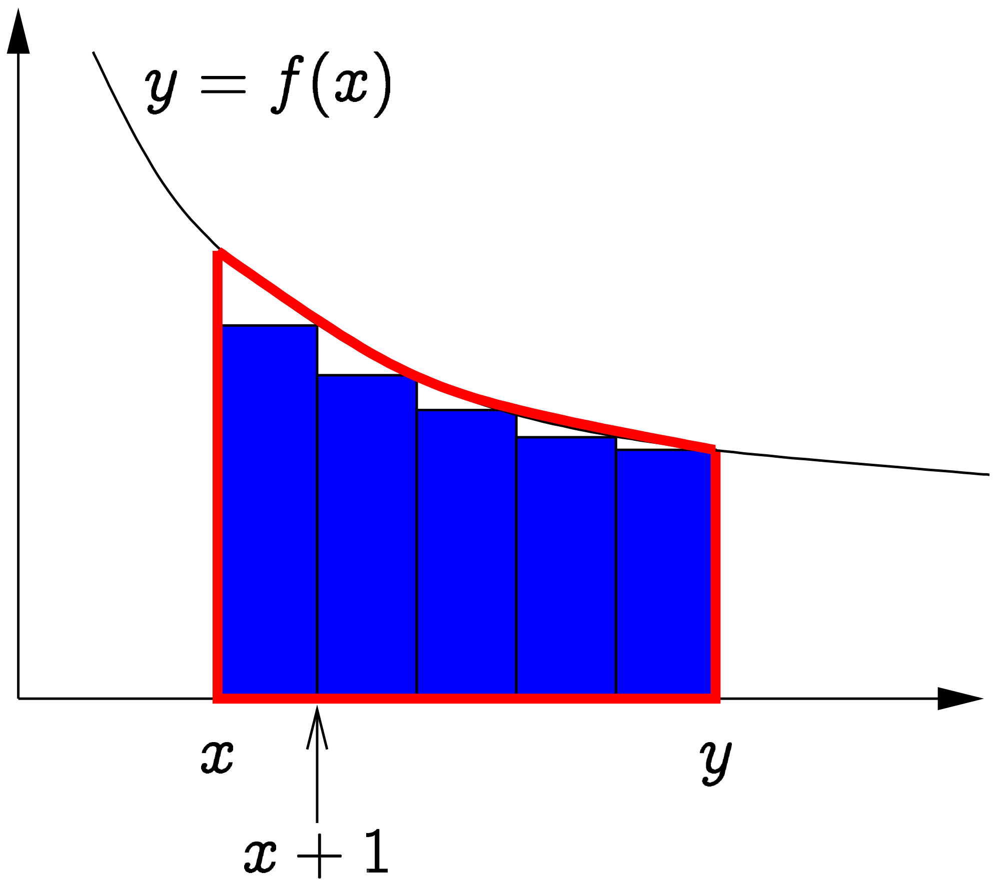
Tõestus. Olgu otsitav võti k ja olgu ta lisatud tabelisse $(i+1)$-ndana. Otsimisel vaadatakse T elemente läbi samas järjekorras kui k-d lisades. Seega on vaadatavate elementide arv sama suur kui ebaõnnestunud otsimisel i kirjega tabelist. See on ülimalt

$$\sum_{i=0}^{n-1} \frac{1}{n} \cdot \frac{1}{1 - i/m} = \frac{m}{n} \sum_{i=0}^{n-1} \frac{1}{m-i} = \frac{1}{\alpha} \sum_{j=m-n+1}^{m} \frac{1}{j} \leqslant$$

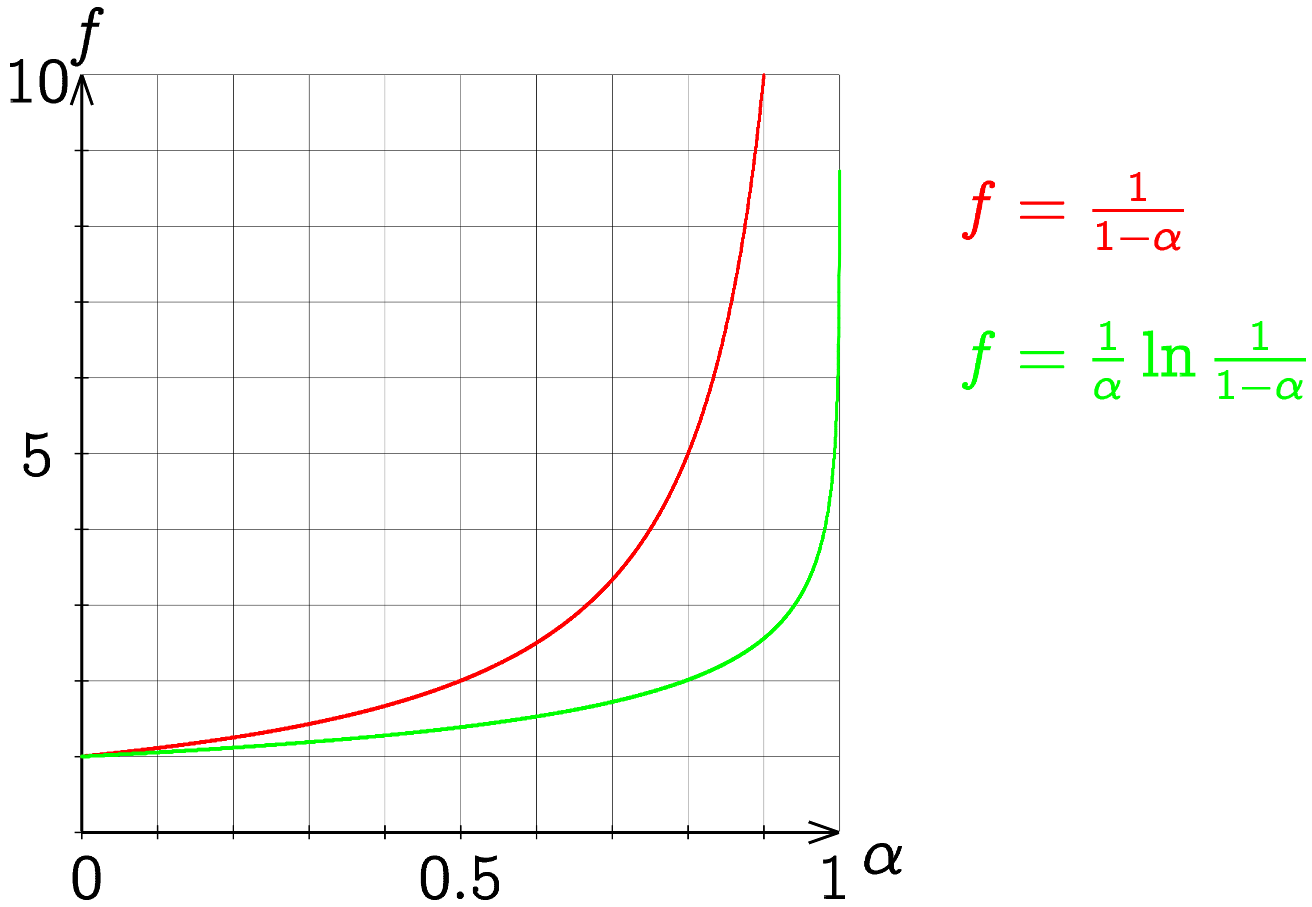
$$\frac{1}{\alpha} \int\limits_{m-n}^{m} \frac{dx}{x} = \frac{1}{\alpha} \left(\ln x \Big|_{m-n}^{m} \right) = \frac{1}{\alpha} \cdot (\ln m - \ln(m-n)) =$$

$$\frac{1}{\alpha} \ln \frac{m}{m-n} = \frac{1}{\alpha} \ln \frac{1}{1-\alpha} \ .$$

□



Kui f on kahanev funktsioon, siis $\sum\limits_{i=x+1}^{y} f(i) \leqslant \int\limits_{x}^{y} f(x)dx$.



f
10
5
0
0.5
1
α
$f = \frac{1}{1-\alpha}$
$f = \frac{1}{\alpha} \ln \frac{1}{1-\alpha}$

Olgu meil antud objektid $x_1, \ldots, x_n$. Me soovime pidada andmestruktuuri, mis vastaks hulkade komplektile $(S_1, \ldots, S_k)$, nii et iga element x_i kuuluks täpselt ühte hulkadest S_j.

Me soovime järgmiseid operatsioone:

- tee_klass(x) lisab juba vaadeldavatele objektidele uue objekti x, samuti lisab ta uue hulka S ja võtab x-i selle ainsaks elemendiks.
- ühenda(x, y) ühendab hulgad, millesse kuuluvad x ja y, üheks hulgaks.
- leia_esindaja(x) tagastab selle hulga S, kuhu x kuulub, mingi elemendi. Seejuures tagastab ta hulga S iga elemendi jaoks sama elemendi.

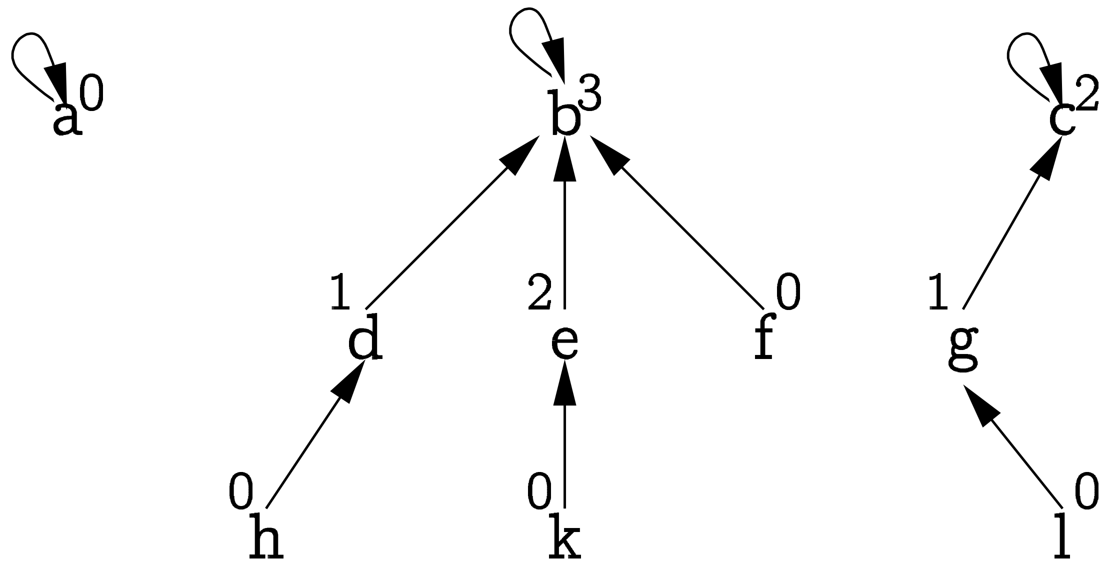
*Galler-Fisher*i meetodil klasside kujutamisel on igal objektil kaks abivälja — $.viit$ ja $.h$.

Klassi kujutatakse puuna tema elementidest, $.viit$ on viit ülemusele. Objekti väli $.h$ on mingi naturaalarv, mis on vähemalt sama suur kui alampuu, mille juureks see objekt on, kõrgus.

Kui mingi objekt on klassi kujutava puu juureks, siis tema $.viit$ viitab talle enesele.

leia_esindaja(x) tagastab x-i sisaldava puu juure.

Näiteks võib ({a}, {b,d,e,f,h,k}, {c,g,l}) kujutatud olla järg-
misel viisil:



tee_klass(x) on

$x.viit := x$; $x.h := 0$

ühenda(x, y) on

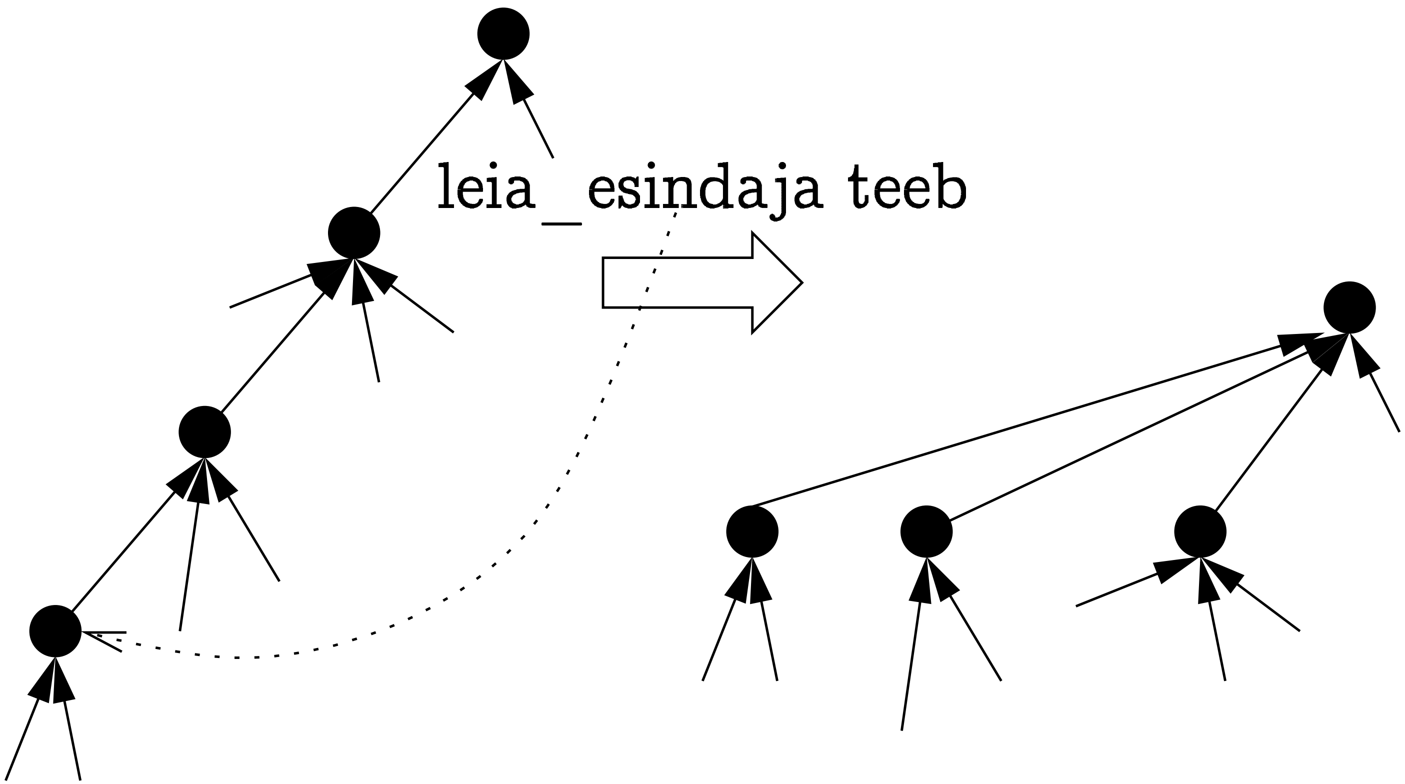
lingi(leia_esindaja(x), leia_esindaja(y))

lingi(x, y) on

if $x.h > y.h$ **then** $y.viit := x$ **else** $x.viit := y$
if $x.h = y.h$ **then** $y.h := y.h + 1$

leia_esindaja(x) on

if $x \neq x.viit$ **then**
    $x.viit :=$ leia_esindaja($x.viit$)
return $x.viit$



Keerukus: kui meil on n objekti ja m operatsiooni tee_klass, ühenda ja leia_esindaja, siis ajakulu on $O(m \log^* n)$.

$\log^* n$ on selline i, et $0 < \underbrace{\log \cdots \log}_{i} n \leqslant 1$.

n-i vahemik	$\log^* n$
$n = 1$	0
$n = 2$	1
$3 \leqslant n \leqslant 4$	2
$5 \leqslant n \leqslant 16$	3
$17 \leqslant n \leqslant 65536$	4
$65537 \leqslant n \leqslant 2^{65536}$	5

Me näitame, et n objekti korral on m operatsiooni tee_klass, lingi ja leia_esindaja ajakulu on $O(m \log^* n)$.

Siis on ka m operatsiooni tee_klass, ühenda ja leia_esindaja ajakulu $O(m \log^* n)$, sest nad sisaldavad ülimalt $3m$ operatsiooni tee_klass, lingi ja leia_esindaja.

Puude struktuuri ja selle muutumise kohta töö jooksul võib öelda järgmist. Olgu x mingi objekt.

- Alguses on x mingi puu juur. Kui ta töö käigus mingil hetkel mittejuureks muutub, siis ei saa ta enam kunagi juureks.
- Kirjutusviis $x.viit = x$ tähendab „x on mingi puu juur“.

Mõned tulemused väljade $.h$ väärtuste kohta ja nende muutumise kohta töö jooksul. Vaatame mingit objekti x.

- Kui $x.viit \neq x$, siis $x.h < x.viit.h$.
- Alguses on $x.h = 0$. Töö jooksul ta ei vähene. Alates hetkest, kus $x.viit \neq x$, $x.h$ enam ei muutu.
- $x.viit.h$ töö jooksul ei vähene.
- Olgu $P(x)$ tippude arv alampuus, mille juureks on x. Siis $P(x) \geqslant 2^{x.h}$.

Tõestused on induktsiooniga üle tehtud operatsioonide arvu.

Kehtigu eelmisel slaidil toodud väited peale mingit arvu operatsioone. Teeme järjekordse operatsiooni.

Kui see oli tee_klass(x), siis saab $x.h$ väärtuseks 0. x on mingi puu juur. $P(x) = 1 = 2^{x.h}$.

Kui see oli lingi(x, y), siis muutused toimuvad ainult x-i ja y-i juures.

Kui enne operatsiooni $x.h \neq y.h$ (üldisust kitsendamata loeme, et $x.h < y.h$), siis pärast operatsiooni on juureks mittevõetud tipu x väli $.h$ väiksem kui juureks võetud tipu y väli $.h$.

Samuti on pärast operatsiooni $P(y) \geqslant 2^{y.h} + 2^{x.h} \geqslant 2^{y.h}$.

Kui enne operatsiooni $x.h = y.h$, siis pärast $y.h = x.h + 1 > x.h$. Samuti $P(y) \geqslant 2^{y.h-1} + 2^{x.h} = 2 \cdot 2^{y.h-1} = 2^{y.h}$.

lingi on ainus operatsioon, mis muudab välja $.h$ väärtusi. Muudetakse tipul, mis on puu juur.

Kui järjekordne operatsioon oli leia_esindaja(x), siis võis $x.viit$ hakata viitama kõrgemale kui varem.

Kuna ennem oli $x.h < x.viit.h$ iga tipu x jaoks, siis puus ülespoole liikudes välja $.h$ väärtused suurenevad.

Välja $x.viit$ muutes siis $x.viit.h$ suurenes.

Järeldus. Tippe, mille välja $.h$ väärtus on vähemalt r, on ülimalt $n/2^r$.

Tõestus. Kui mingi tipu x jaoks omistatakse $x.h := r$, siis peab puus, mille juureks on x, olema vähemalt 2^r tippu. Nende tippude (v.a. x ise) aste on väiksem kui r ja enam ei muutu.

Kui mõne teise tipu y jaoks omistatakse $y.h := r$, siis peab puus, mille juures on y, olema vähemalt 2^r tippu. Need tipud peavad olema erinevad puu, mille juureks on x, tippudest, sest x ei kuulu puusse, mille juureks on y.

Seega saab ülimalt iga 2^r-s tipp saada $.h$ väärtuseks r või enam. Muuhulgas on siis iga tipu välja $.h$ väärtus $\leqslant \log n$.

Üks operatsioon tee_klass või lingi vajab konstantset aega.

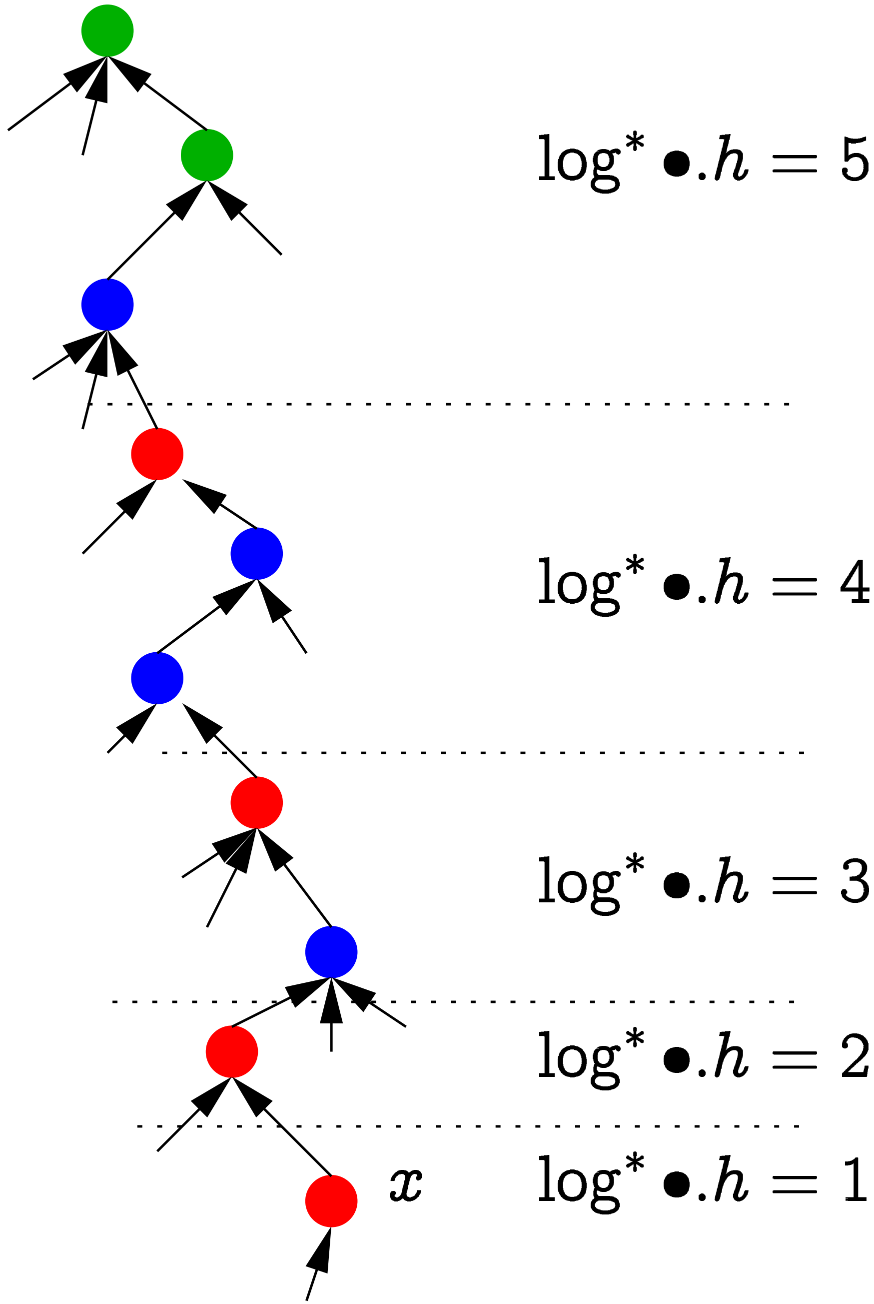
Ühe operatsiooni leia_esindaja(x) tööaeg on proportsionaalne objekti x kaugusega teda sisaldava puu juurest selle operatsiooni tegemise hetkel.

Me näitame, et peale m-i operatsiooni on nende kauguste summa $O(m \log^* n)$.

Vaatame mingit operatsiooni leia_esindaja(x).

Vaatame ahelat tipust x teda sisaldava puu juureni. Värvime sellel ahelal tipud järgmiselt:

- Juure ja tema vahetu järglase värvime roheliseks.
- Mõne teise tipu v sellelt ahelalt värvime punaseks parajasti siis, kui $\log^* y.h \neq \log^* y.viit.h$.
  - Defineerime $\log^* 0 := -1$.
- Ülejäänud tipud värvime siniseks.
  - Neil tippudel $\log^* y.h = \log^* y.viit.h$.


$\log^* \bullet.h = 5$
$\log^* \bullet.h = 4$
$\log^* \bullet.h = 3$
$\log^* \bullet.h = 2$
x
$\log^* \bullet.h = 1$

Igal operatsioonil leia_esindaja on ülimalt kaks rohelist ja ülimalt $\log^* \log n = \log^* n - 1$ punast tippu.

Kui mingi tipp x on kunagi olnud punane, siis ei saa ta enam kunagi edaspidi olla sinine, sest

- $x.h$ enam ei muutu, seega ei muutu ka $\log^* x.h$;
- $\log^* x.h < \log^* x.viit.h$;
- $x.viit.h$ saab üksnes suureneda, seega ka $\log^* x.viit.h$ saab üksnes suureneda.

Uurime, mitu korda saab üks tipp olla sinine.

Kui tipp x on mingis operatsioonis leia_esindaja sinine, siis muutub tema otsene ülemus. S.t. $x.viit.h$ suureneb selle operatsiooni käigus.

Ta saab suureneda ainult senikaua kuni $\log^* x.viit.h = \log^* x.h$, muidu muutub tipp punaseks.

Olgu $S(x)$ kordade arv, mil tipp x võib olla sinine. Siis on m operatsiooni tehes operatsioonidel leia_esindaja vaadeldavate ahelate pikkuste summa ülimalt

$$2m + m(\log^* n - 1) + \sum_{i=1}^{n} S(x_i) \quad .$$

Tähistame: $\mathbf{2}^{*-2} := -1$, $\mathbf{2}^{*-1} := 0$, $\mathbf{2}^{*0} := 1$, $\mathbf{2}^{*i} := 2^{\mathbf{2}^{*i-1}}$.

S.t. $\mathbf{2}^{*i} = \left.2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}\right\} i$

Siis $\log^* n = i$ parasjagu siis, kui $\mathbf{2}^{*i-1} + 1 \leqslant n \leqslant \mathbf{2}^{*i}$.

Olgu $j = \log^* x.h$. Siis $x.viit.h$ võib suureneda ülimalt $\mathbf{2}^{*j} - \mathbf{2}^{*j-1} - 1$ korral, nii et $\log^* x.viit.h$ ei saa suuremaks kui $\log^* x.h$.

S.t. $S(x) \leqslant \mathbf{2}^{*\log^* x.h} - \mathbf{2}^{*(\log^* x.h)-1} - 1$.

Olgu $N(j) = |\{x : \log^* x.h = j\}|$. Siis on m operatsiooni tehes operatsioonidel leia_esindaja vaadeldavate ahelate pikkuste summa ülimalt

$$2m + m(\log^* n - 1) + \sum_{j=-1}^{\log^* n - 1} N(j)(\mathbf{2}^{*j} - \mathbf{2}^{*j-1} - 1)\ .$$

Peale selle,

$$N(j) \leqslant \sum_{r=\mathbf{2}^{*j-1}+1}^{\mathbf{2}^{*j}} \frac{n}{2^r},$$

sest tippe, mille välja $.h$ väärtus on vähemalt r, on ülimalt $n/2^r$.

$$
\begin{aligned}
N(j) &\leqslant \sum_{r=\mathbf{2}^{*j-1}+1}^{\mathbf{2}^{*j}} \frac{n}{2^r} = \frac{n}{2^{\mathbf{2}^{*j-1}+1}} \sum_{r=0}^{\mathbf{2}^{*j}-\mathbf{2}^{*j-1}-1} \frac{1}{2^r} \\
&\leqslant \frac{n}{2^{\mathbf{2}^{*j-1}+1}} \sum_{r=0}^{\infty} \frac{1}{2^r} = \frac{n \cdot 2}{2^{\mathbf{2}^{*j-1}+1}} = \frac{n}{2^{\mathbf{2}^{*j-1}}} \\
&= \begin{cases} 2n, & \text{kui } j = -1 \\ \dfrac{n}{\mathbf{2}^{*j}}, & \text{kui } j \geqslant 0 \end{cases}
\end{aligned}
$$

$$2m + m(\log^* n - 1) + \sum_{j=-1}^{\log^* n - 1} N(j)(\mathbf{2}^{*j} - \mathbf{2}^{*j-1} - 1) \leqslant$$

$$m(\log^* n + 1) + \sum_{j=-1}^{\log^* n - 1} N(j)\mathbf{2}^{*j} \leqslant$$

$$m(\log^* n + 1) + 2n\mathbf{2}^{*-1} + \sum_{j=0}^{\log^* n - 1} \frac{n}{\mathbf{2}^{*j}} \mathbf{2}^{*j} =$$

$$m(\log^* n + 1) + \sum_{j=0}^{\log^* n - 1} n =$$

$$m(\log^* n + 1) + n \log^* n = O(m \log^* n)$$