Biltré

Bergur Snorrason

February 19, 2024

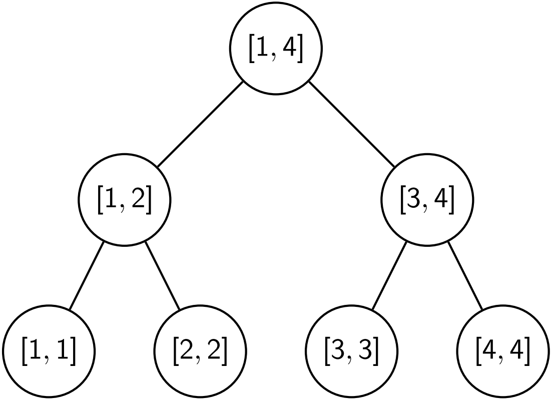
Dæmi

- Gefinn er listi með n tölum.
- Næst koma q fyrirspurnir, þar sem hver er af einni af tveimur gerðum:
  - Bættu k við i-tu töluna.
  - Reiknaðu summu allra talna á bilinu $[i, j]$.
- Einföld útfærlsa á þessum fyrirspurnum gefur okkur $\mathcal{O}(1)$ fyrir þá fyrri og $\mathcal{O}(n)$ fyrir þá seinni.
- Þar sem allar (eða langflestar) fyrirspurnir gætu verið af seinni gerðin yrði lausnin í heildina $\mathcal{O}(qn)$.
- Það er þó hægt að leysa þetta dæmi hraðar.
- Algengt er að nota til þess *biltré* (e. *segment tree*).

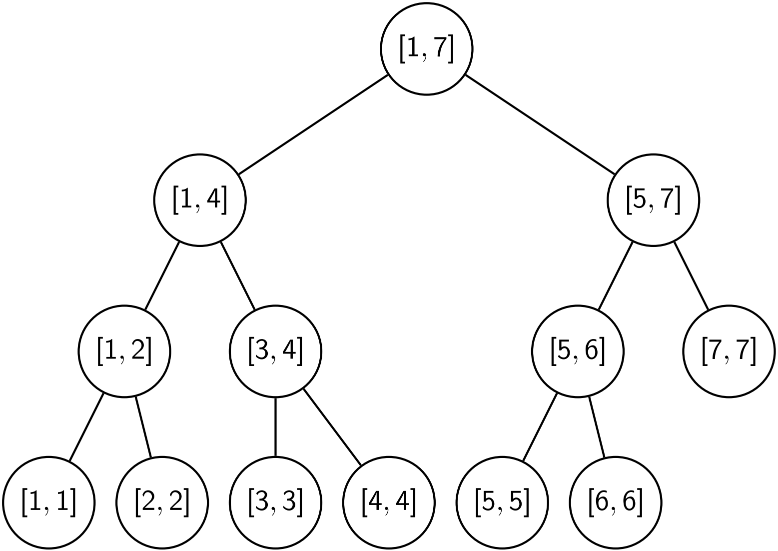
Biltré

- Biltré er tvíundartré sem geymir svör við vissum fyrirspurnum af seinni gerðinni.
- Rótin geymir svar við fyrirspurninni `1 n` og ef hnútur geymir svarið við `i j` þá geyma börn hans svör við `i m` og `(m + 1) j`, þar sem m er miðja heiltölubilsins $[i, j]$.
- Þeir hnútar sem geyma svar við fyrirspurnum af gerðinni `i i` eru lauf trésins.
- Takið eftir að laufin geyma þá gildin í listanum og aðrir hnútar geyma summu barna sinna.
- Þegar við útfærum tréð geymum við það eins og hrúgu.

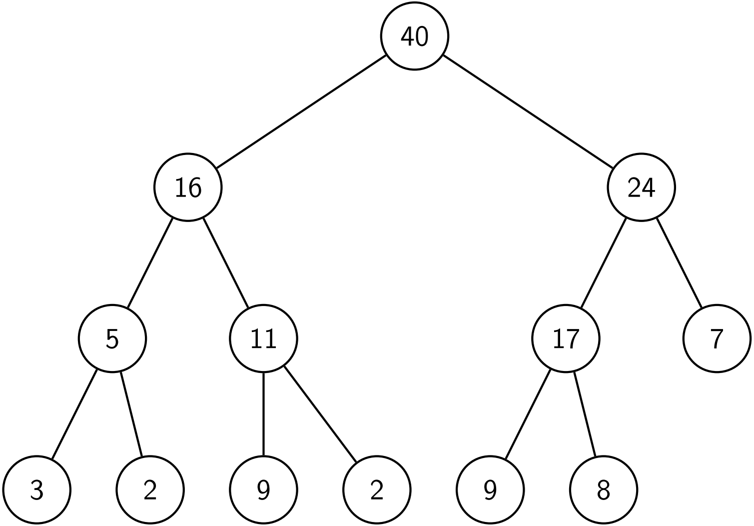
Mynd af biltré, $n = 4$



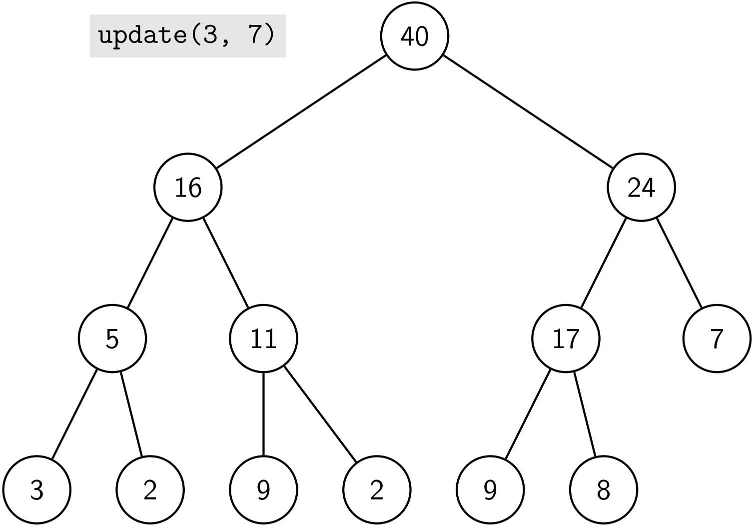
Mynd af biltré, $n = 7$



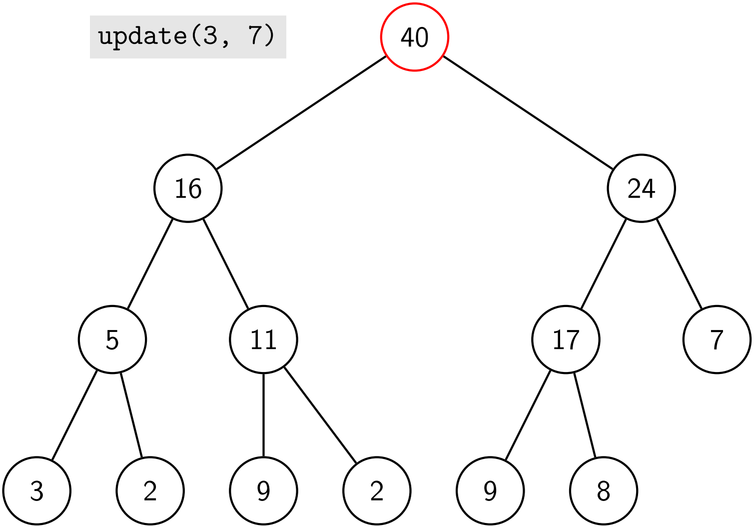
- Gerum ráð fyrir að við höfum biltré eins og lýst er að ofan og látum H tákna hæð trésins.
- Hvernig getum við leyst fyrirspurnirnar á glærunni á undan, og hver er tímaflækjan?
- Fyrri fyrirspurnin er einföld.
- Ef við eigum að bæta k við i-ta stakið finnum við fyrst laufið sem svarar til fyrirspurnar `i i`, bætum k við gildið þar og förum svo upp í rót í gegnum foreldrin og uppfærum á leiðinni gildin í þeim hnútum sem við lendum.
- Þar sem við heimsækjum bara þá hnúta sem eru á veginum frá rót til laufs (mest H hnútar) er tímaflækjan á fyrri fyrirspurninni $\mathcal{O}(H)$.


40
16
24
5
11
17
7
3
2
9
2
9
8

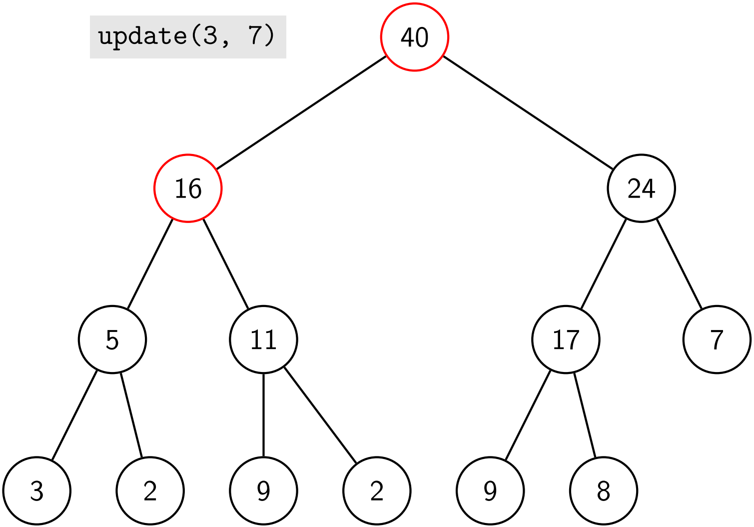
`update(3, 7)`

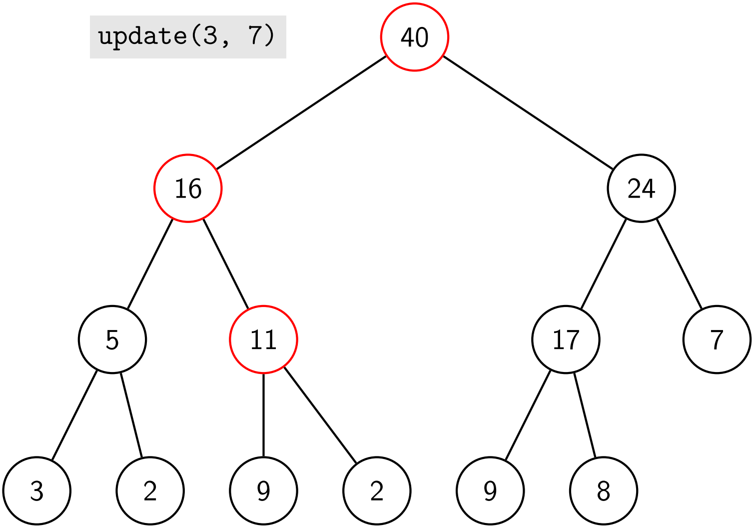


`update(3, 7)`

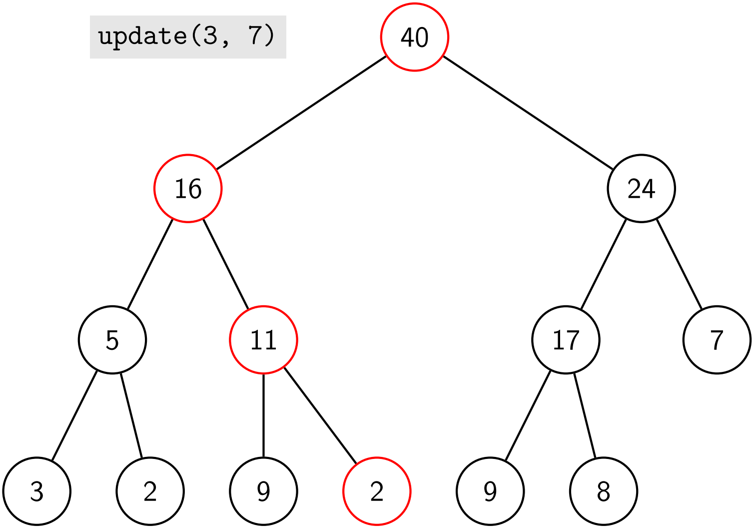


`update(3, 7)`

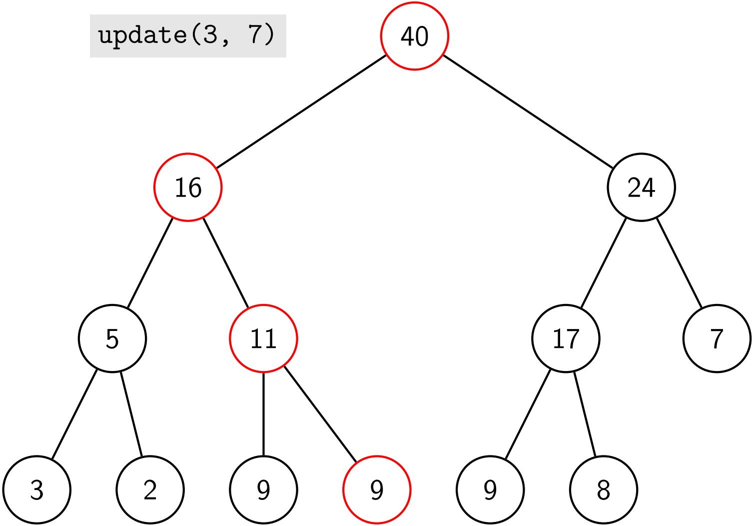


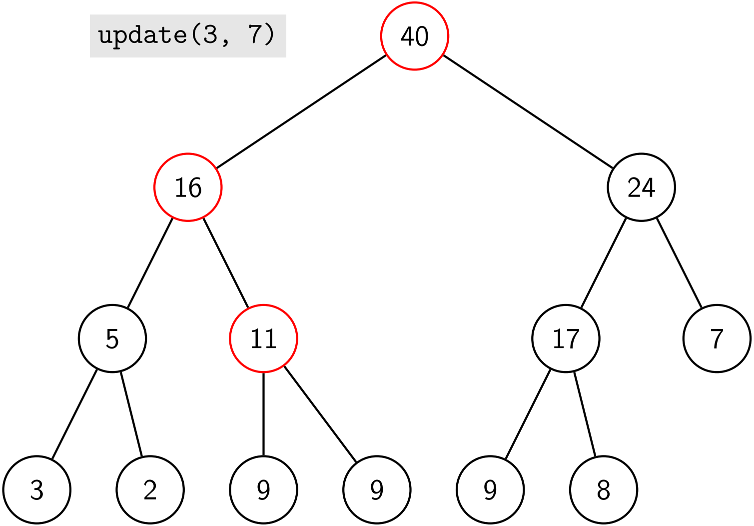

update(3, 7)
40
16
24
5
11
17
7
3
2
9
2
9
8

`update(3, 7)`

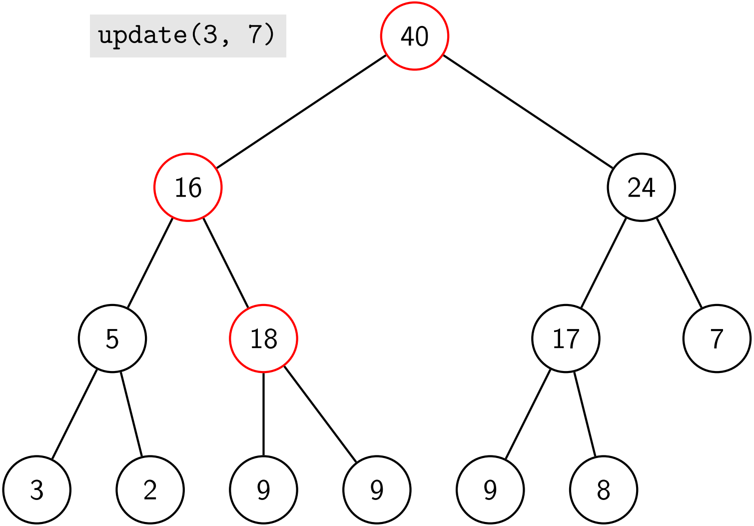


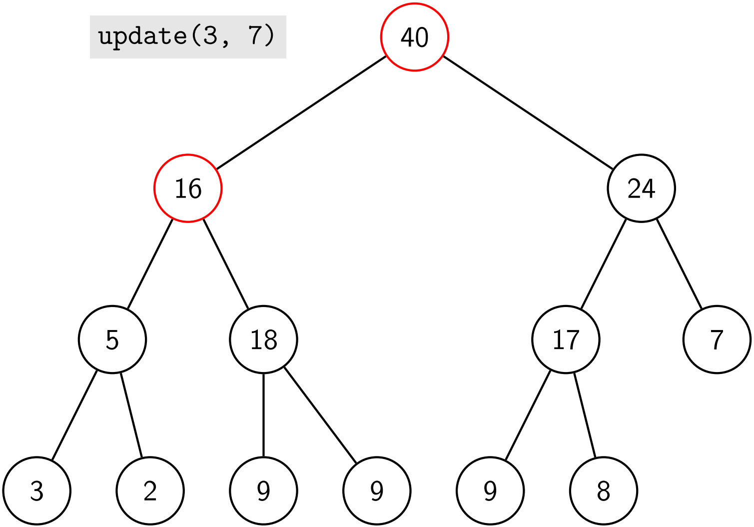
`update(3, 7)`



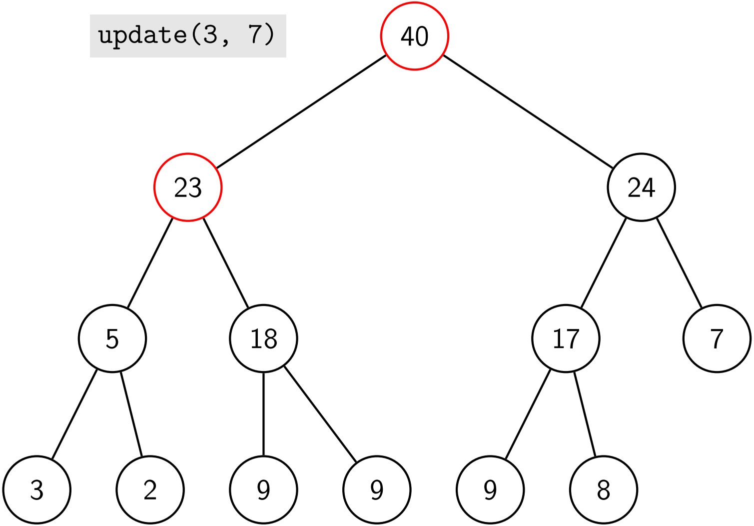

update(3, 7)
40
16
24
5
11
17
7
3
2
9
9
9
8

`update(3, 7)`

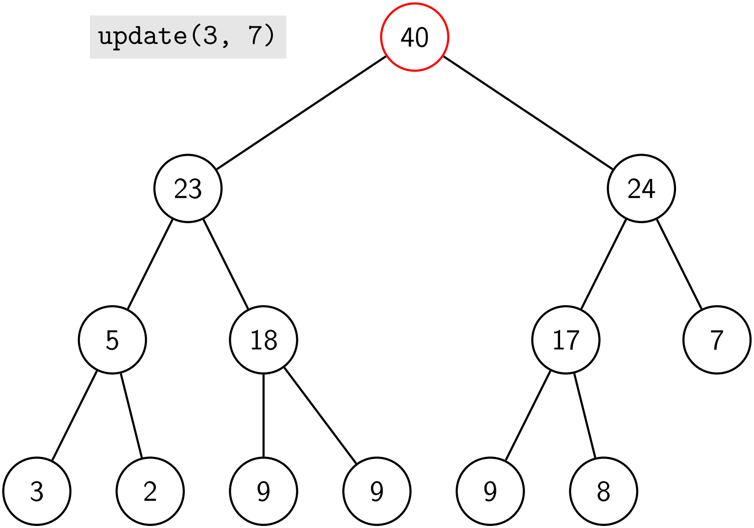



update(3, 7)
40
16
24
5
18
17
7
3
2
9
9
9
8

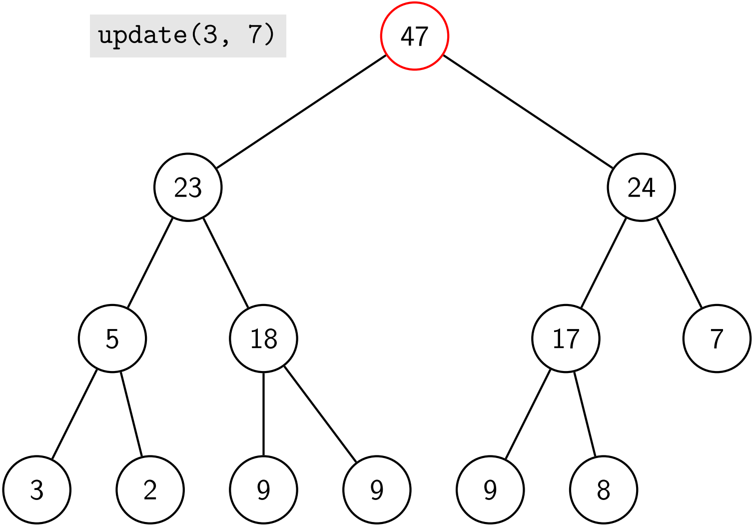
`update(3, 7)`

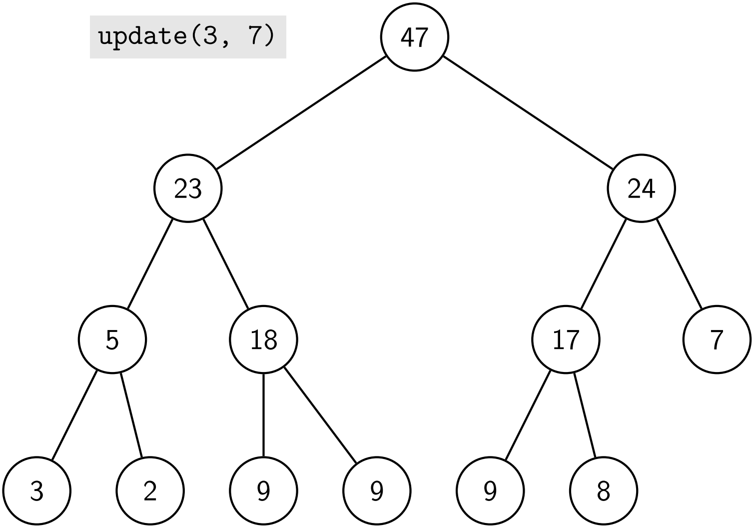


`update(3, 7)`

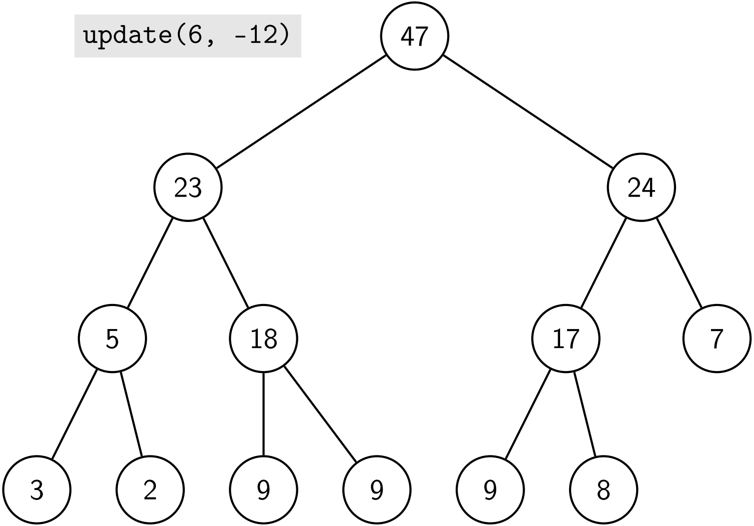


`update(3, 7)`

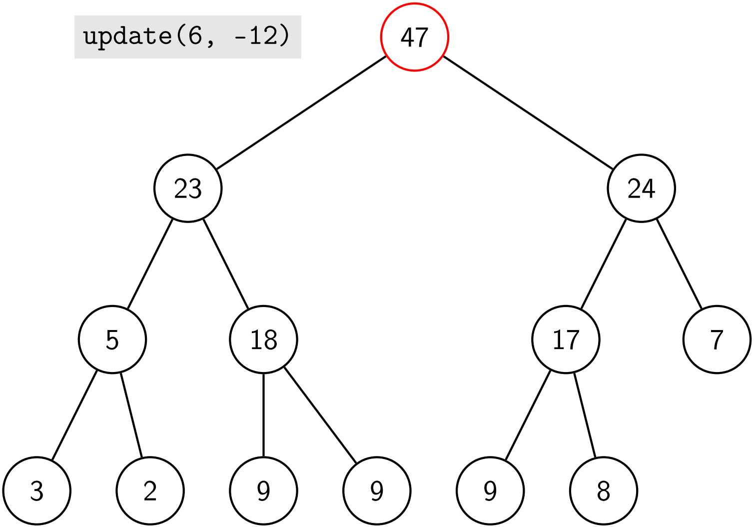



update(3, 7)
47
23
24
5
18
17
7
3
2
9
9
9
8

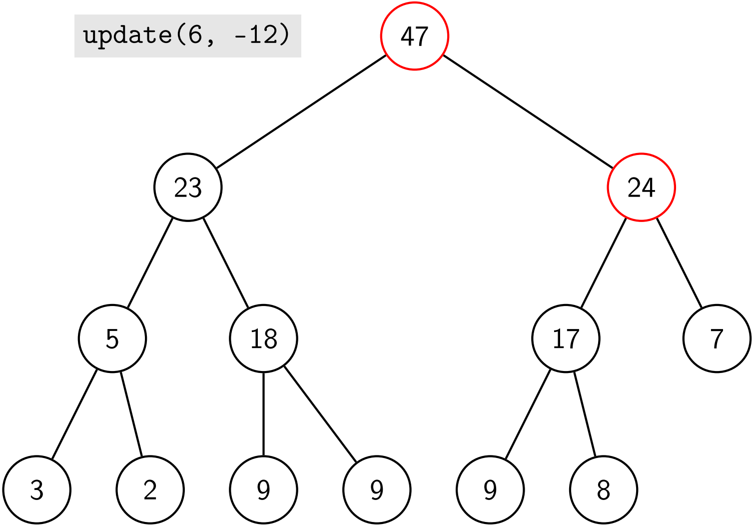
`update(6, -12)`



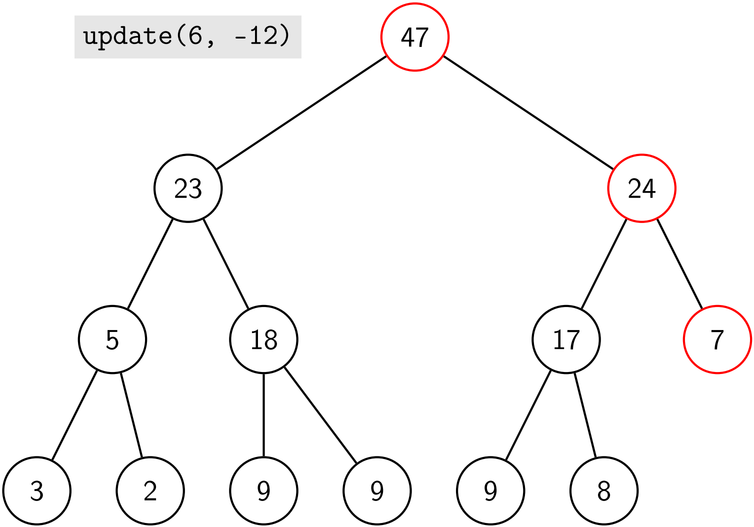
`update(6, -12)`



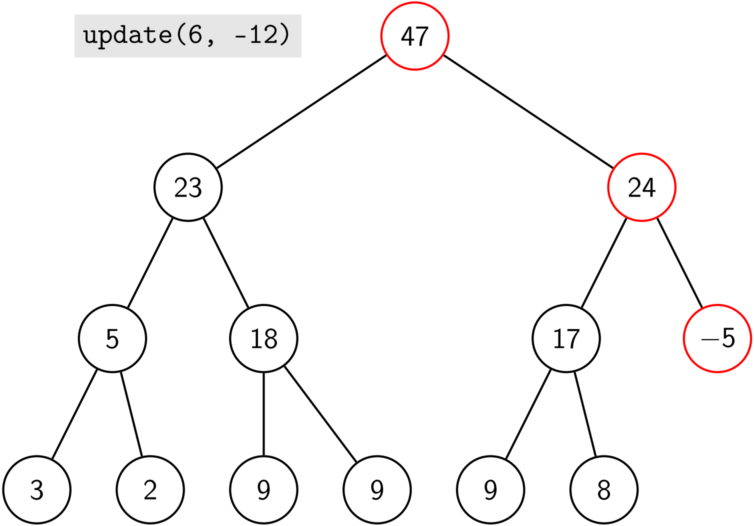
`update(6, -12)`



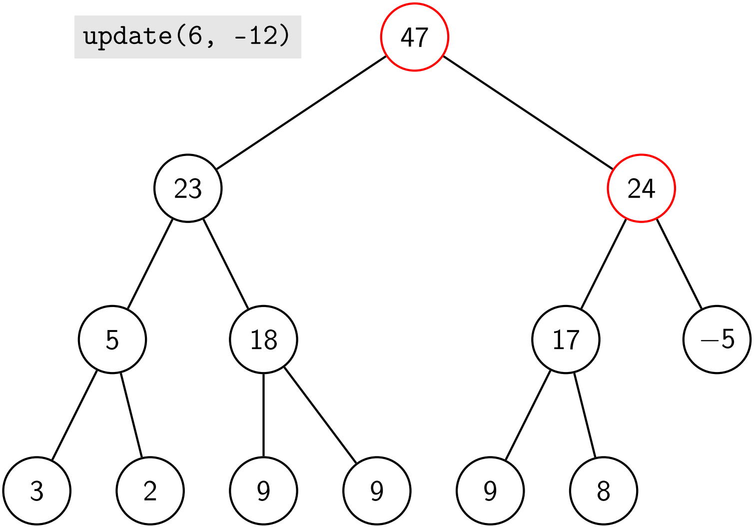
`update(6, -12)`



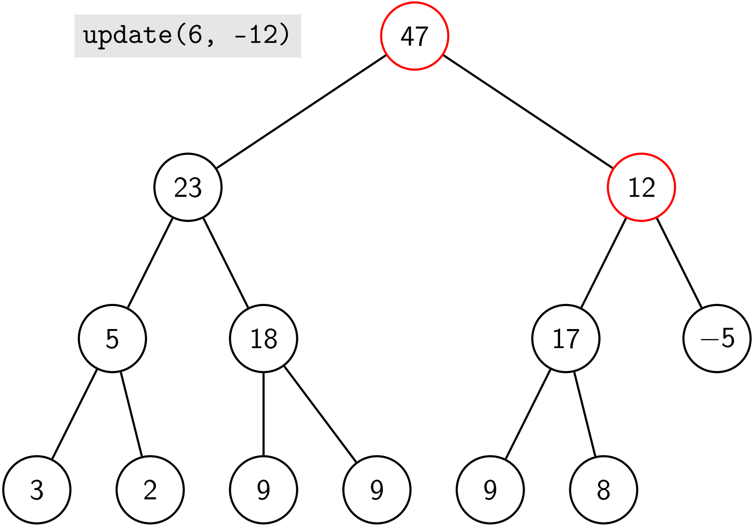
`update(6, -12)`



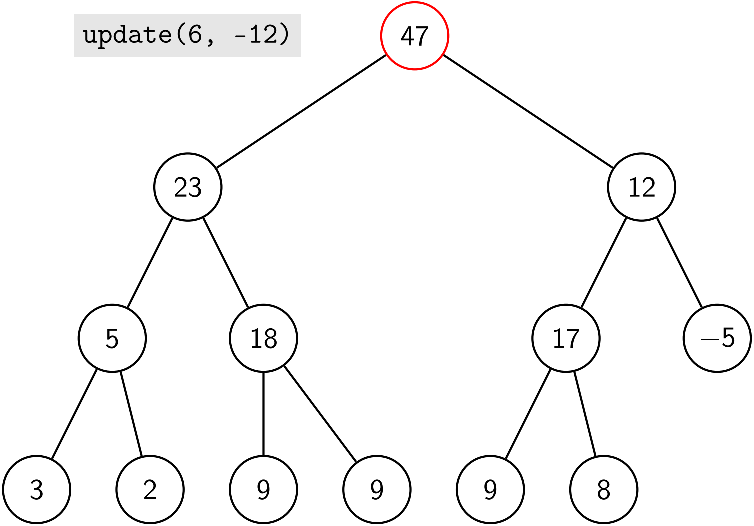
`update(6, -12)`



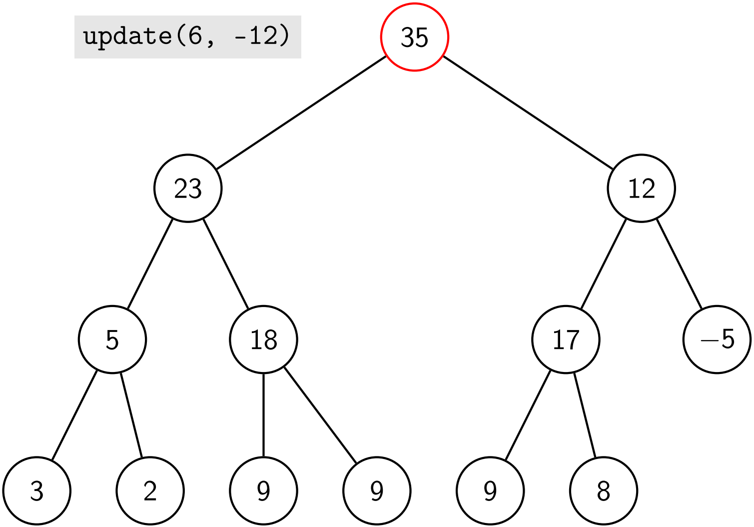
`update(6, -12)`



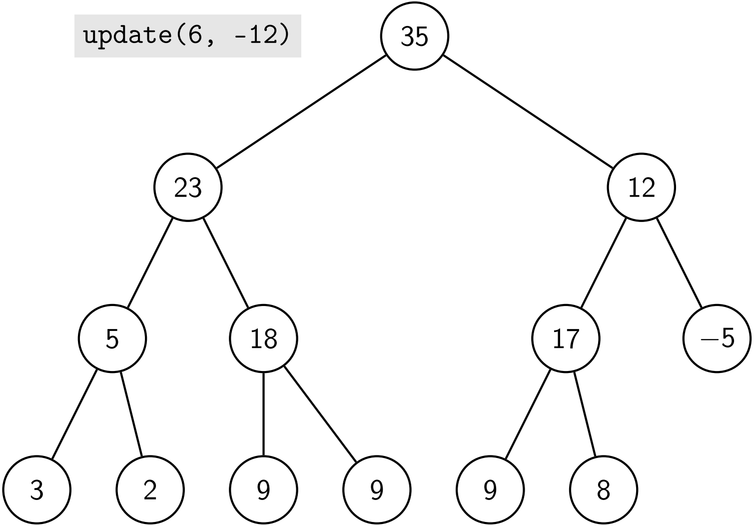
`update(6, -12)`

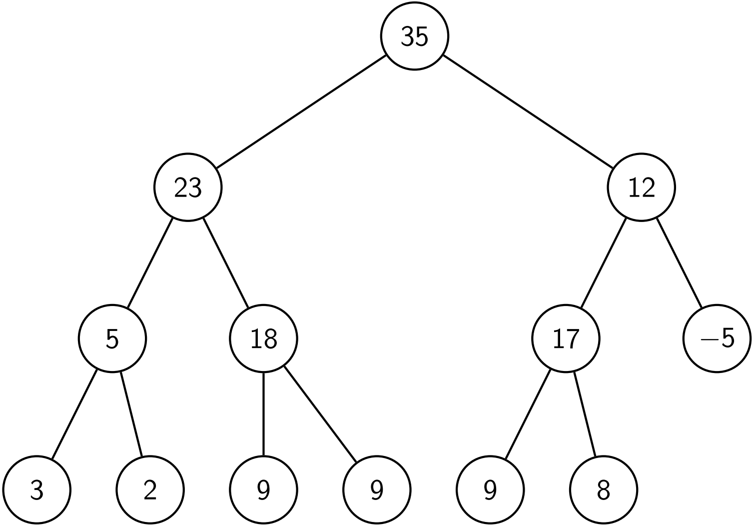


`update(6, -12)`



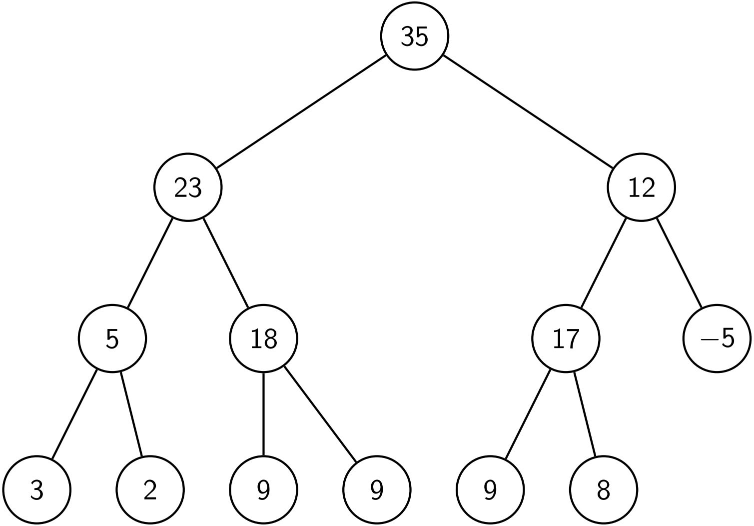
`update(6, -12)`



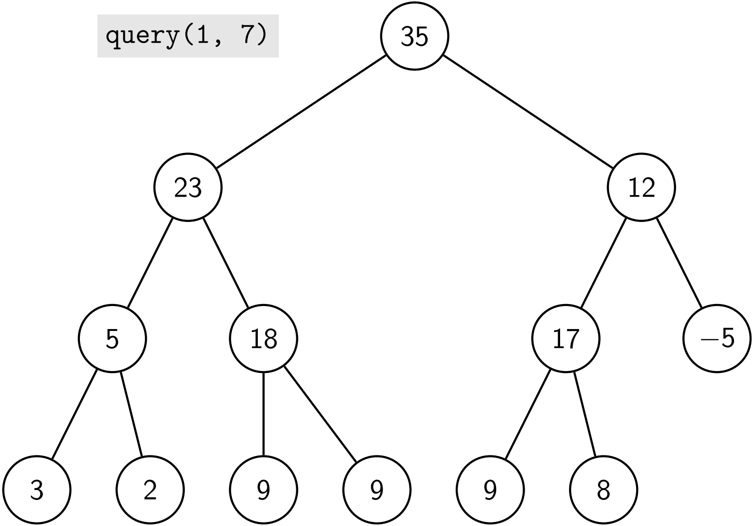

35
23
12
5
18
17
−5
3
2
9
9
9
8

```
void urec(int i, int j, int x, int y, int e)
{
    if (i == j) p[e] += y;
    else
    {
        int m = (i + j)/2;
        if (x <= m) urec(i, m, x, y, LEFT(e));
        else urec(m + 1, j, x, y, RIGHT(e));
        p[e] = p[LEFT(e)] + p[RIGHT(e)];
    }
}
void update(int x, int y)
{
    return urec(0, p[0] - 1, x, y, 1);
}
```

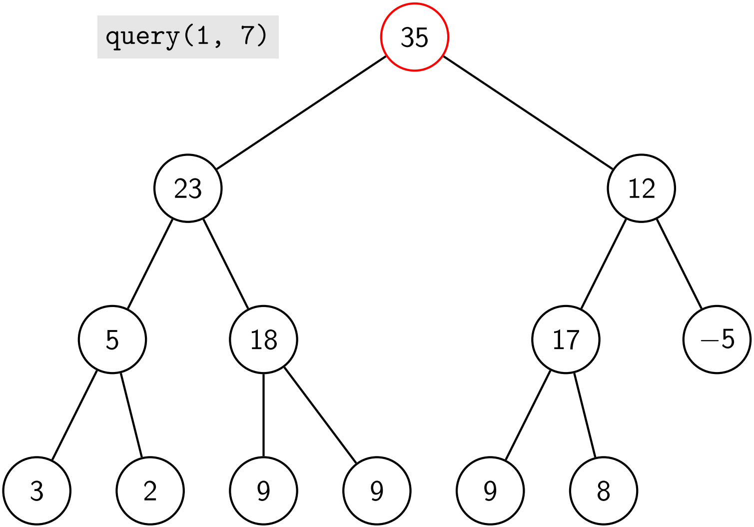
- Seinni fyrirspurnin er ögn flóknari.
- Auðveldast er að ímynda sér að við förum niður tréð og leitum að hvorum endapunktinum fyrir sig.
- Á leiðinni upp getum við svo púslað saman svarinu, eftir því hvort við erum að skoða hægri eða vinstri endapunktinn.
- Til dæmis, ef við erum að leita að hægri endapunkti x og komum upp í bil $[i, j]$ þá bætum við gildinu í hnút $[i, m]$ við það sem við höfum reiknað hingað til ef $x \in [m+1, j]$, en annars bætum við engu við (því x er hægri endapunkturinn).
- Við göngum svona upp þar til við lendum í bili sem inniheldur hinn endapunktinn.
- Með sömu rökum og áðan er tímaflækjan $\mathcal{O}(H)$.


35
23
12
5
18
17
−5
3
2
9
9
9
8

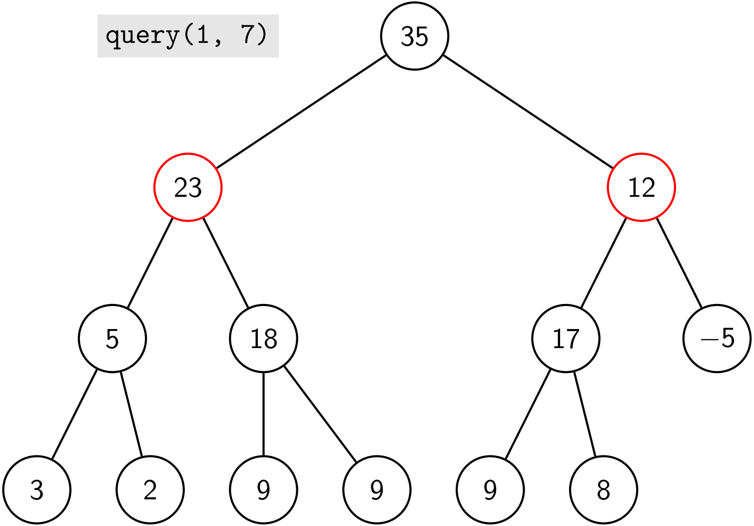
`query(1, 7)`



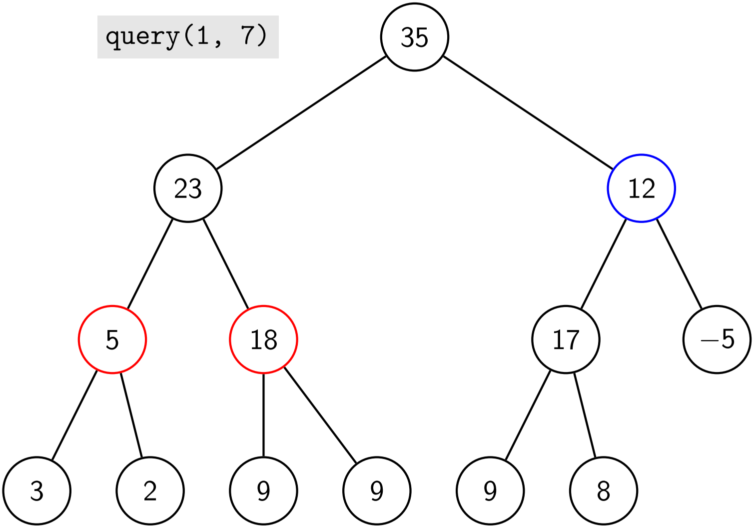
`query(1, 7)`



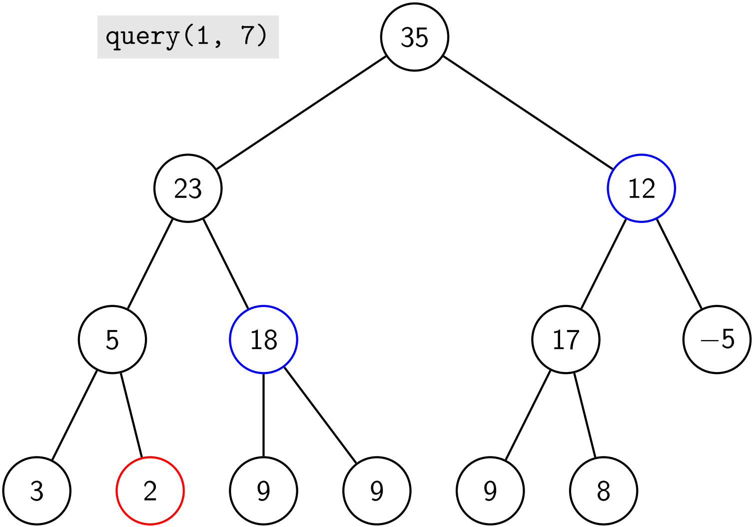
`query(1, 7)`



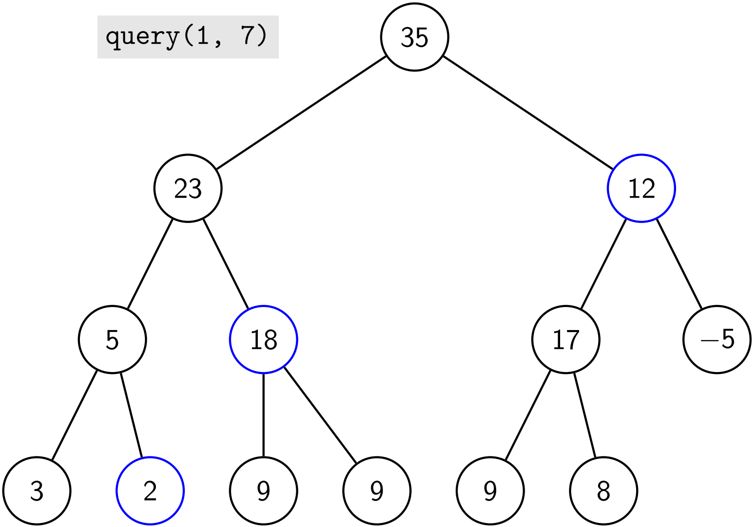
`query(1, 7)`



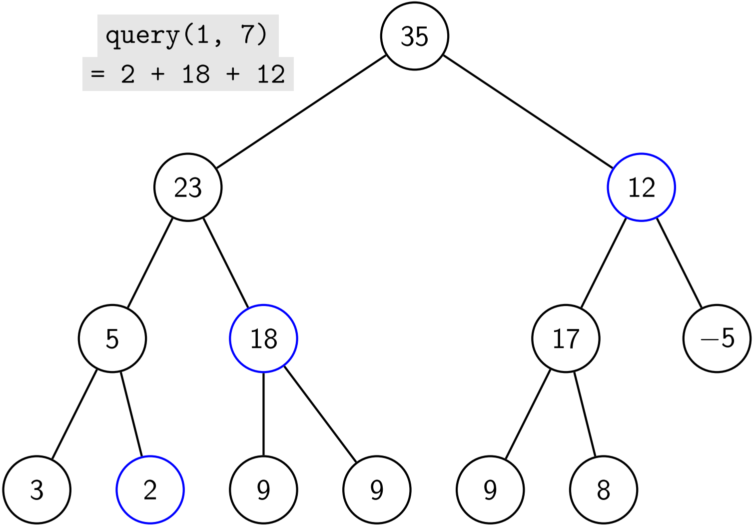
`query(1, 7)`



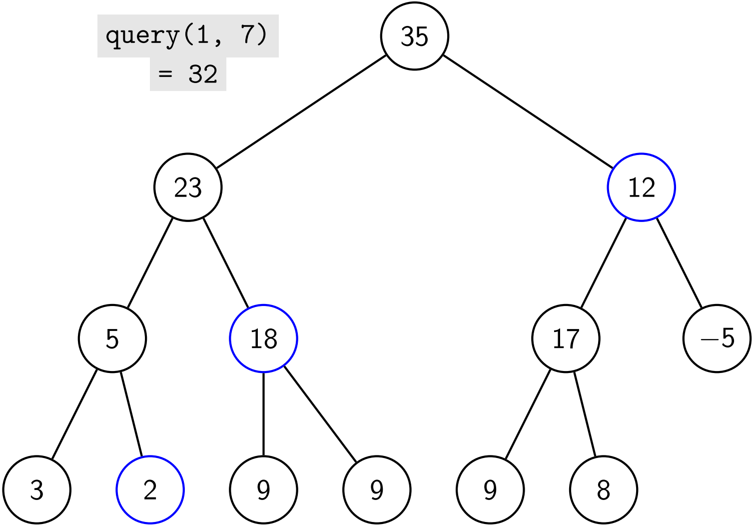
`query(1, 7)`



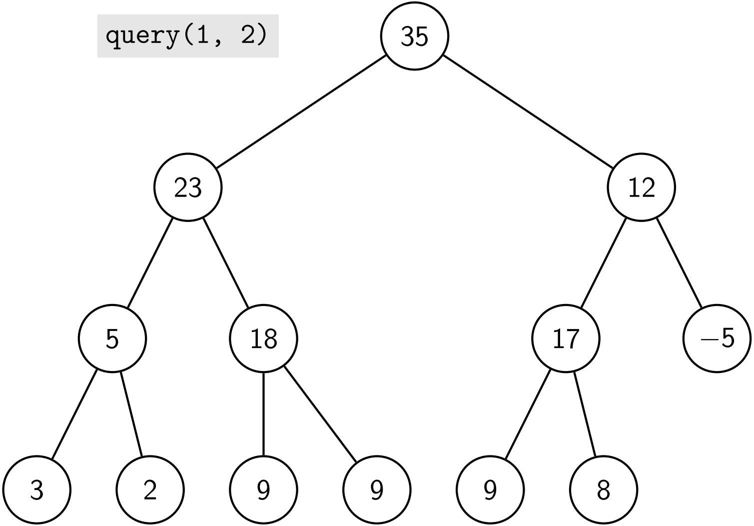
```
query(1, 7)
= 2 + 18 + 12
```



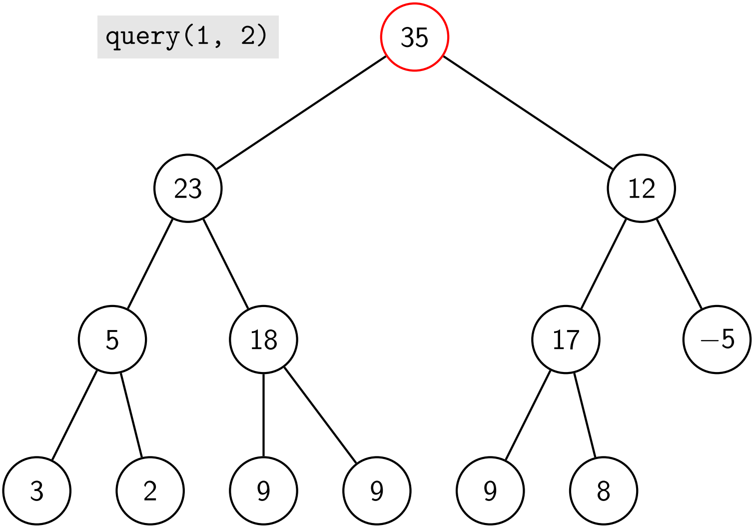
`query(1, 7)`
`= 32`



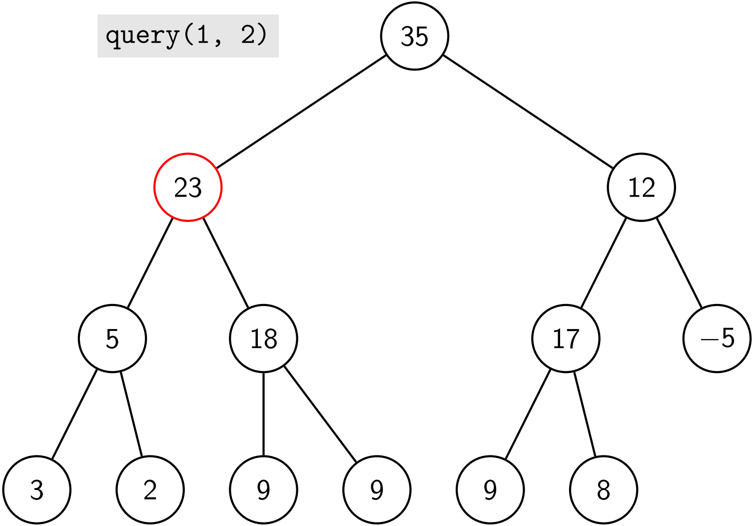
`query(1, 2)`



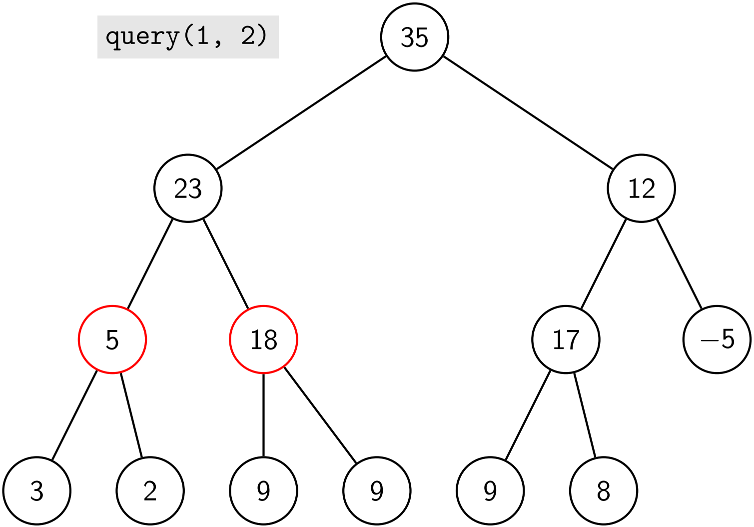
`query(1, 2)`

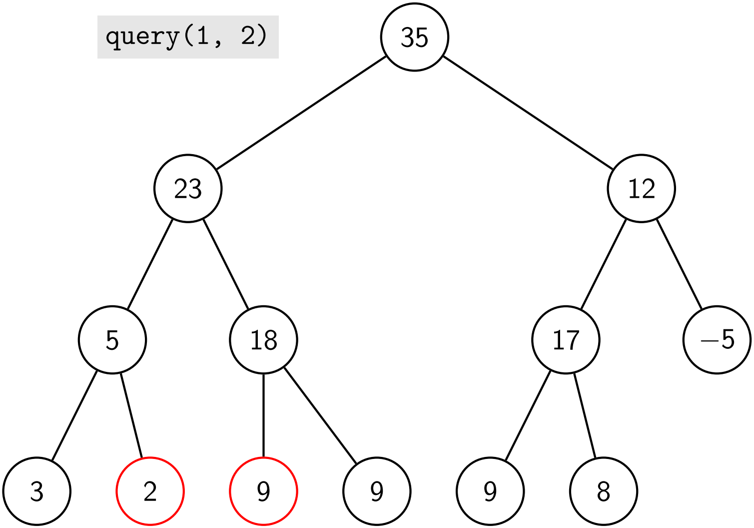


`query(1, 2)`

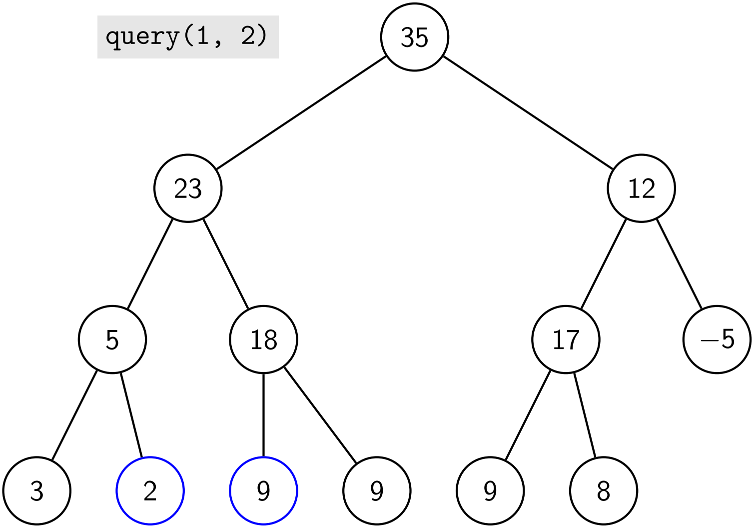


`query(1, 2)`

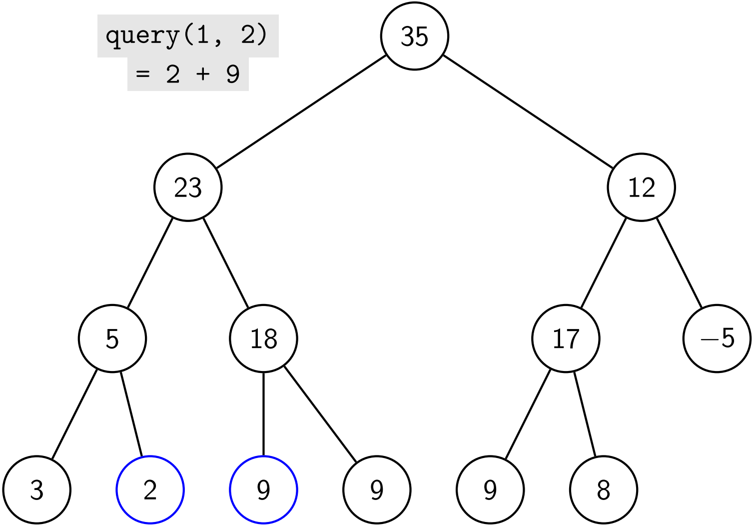



query(1, 2)
35
23
12
5
18
17
−5
3
2
9
9
9
8

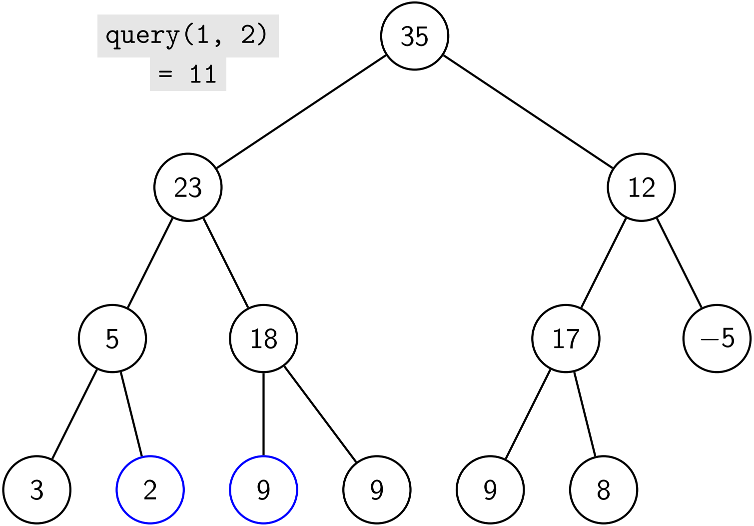
`query(1, 2)`

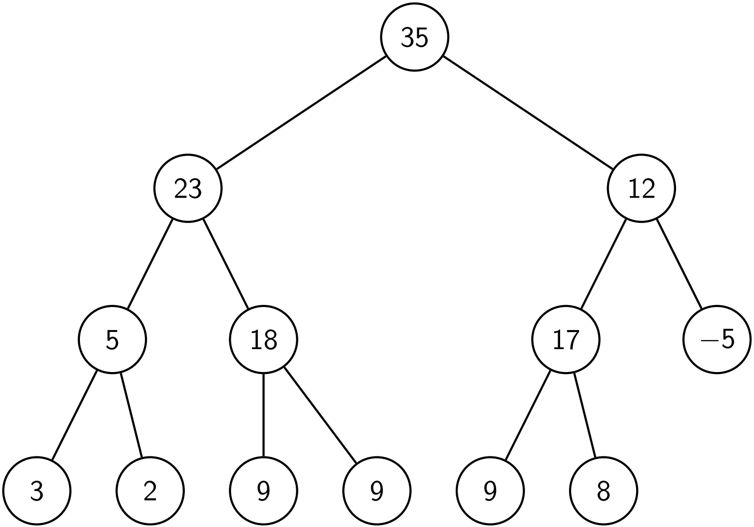


`query(1, 2)`
`= 2 + 9`



```
query(1, 2)
= 11
```




35
23
12
5
18
17
−5
3
2
9
9
9
8

Biltré í `C`

```c
int qrec(int i, int j, int x, int y, int e)
{
    if (x == i && y == j) return p[e];
    int m = (i + j)/2;
    if (y <= m) return qrec(i, m, x, y, LEFT(e));
    if (x > m) return qrec(m + 1, j, x, y, RIGHT(e));
    return qrec(i, m, x, m, LEFT(e)) + qrec(m + 1, j, m + 1, y, RIGHT(e));
}
int query(int x, int y)
{
    return qrec(0, p[0] - 1, x, y, 1);
}
```

Tímaflækja biltrjáa

- Þar sem lengd hvers bils sem hnútur svarar til helmingast þegar farið er niður tréð er $\mathcal{O}(H) = \mathcal{O}(\log n)$.
- Við erum því komin með lausn á upprunalega dæminu sem er $\mathcal{O}(q \cdot \log n)$.
- Þetta væri nógu hratt ef, til dæmis, $n = q = 10^6$.
- Tökum annað dæmi.

Annað dæmi

- Fyrsta lína inntaksins inniheldur tvær jákvæðar heiltölur, n og m, minni en 10^5.
- Næsta lína inniheldur n heiltölur, á milli -10^9 og 10^9.
- Næstu m línur innihalda fyrirspurnir, af tveimur gerðum.
- Fyrri gerðin hefst á `1` og inniheldur svo tvær tölur, x og y. Hér á að setja x-tu töluna sem y.
- Seinni gerðin hefst á `2` og inniheldur svo tvær tölu, x og y. Hér á að prenta út stærstu töluna á hlutbilinu $[x, y]$ í talnalistanum.
- Hvernig leysum við þetta?

- Við getum leyst þetta með biltrjám.
- Í stað þess að láta hnúta (sem eru ekki lauf) geyma summu barna sinna, þá geyma þeir stærra stak barna sinna.

Lausn

```c
int qrec(int i, int j, int x, int y, int e)
{
    if (x == i && y == j) return p[e];
    int m = (i + j)/2;
    if (y <= m) return qrec(i, m, x, y, LEFT(e));
    if (x > m) return qrec(m + 1, j, x, y, RIGHT(e));
    return max(qrec(i, m, x, m, LEFT(e)), qrec(m + 1, j, m + 1, y, RIGHT(e)));
}
int query(int x, int y)
{
    return qrec(0, p[0] - 1, x, y, 1);
}

void urec(int i, int j, int x, int y, int e)
{
    if (i == j) p[e] = y;
    else
    {
        int m = (i + j)/2;
        if (x <= m) urec(i, m, x, y, LEFT(e));
        else urec(m + 1, j, x, y, RIGHT(e));
        p[e] = max(p[LEFT(e)], p[RIGHT(e)]);
    }
}
void update(int x, int y)
{
    return urec(0, p[0] - 1, x, y, 1);
}
```

- Leysa má ýmis dæmi af þessari gerð, með biltrjám.
- Þessi dæmi eru yfirleitt kölluð *punkt-uppfærslur, bil-fyrirspurnir* (e. *point-update, range-query*).
- Algengt er að sýna næst hvernig nota megi biltré til að leysa *bil-uppfærslur, punkt-fyrirspurnir* (e. *range-update, point-query*).
- Þetta er, í grófum dráttum, gert með því að snúa trjánum við.
- Við munum ekki skoða þetta.
- Við tökum frekar fyrir *lygn biltré*.
- Þau leyfa okkur að leysa *bil-uppfærslur, bil-fyrirspurnir* (e. *range-update, range-query*).

Lygn dreifing

- Sem beinagrind munum við nota biltrjáa útfærsluna sem við notuðum til að leysa fyrsta dæmið.
- Við munum nú láta fyrri fyrirspurnina, `i j k`, þýða "Bættu k við allar tölur á bilinu $[i, j]$".
- Uppfærslan er framkvæmd á svipaðan hátt og fyrirspurnirnar eru.
- Við geymum í öðrum tréi þær uppfærslur sem við eigum eftir að framkvæma.
- Í hverri endurkvæmni (bæði uppfærslum og fyrirspurnum) dreifum við uppfærslunum í hnútnum niður á við.
- Þetta kallast *lygn dreifing* (e. *lazy propagation*), því við framkvæmum hana bara þegar nauðsyn krefur.
- Ef biltré hefur lygna dreifingu köllum við það *lygnt biltré* (e. *segment tree with lazy propagation*).

- Látum $i < j$ vera heiltölur þannig að bilið $[i, j]$ svara til hnúts í biltréi og m vera miðpunkt heiltölu bilsins $[i, j]$.
- Gerum ráð fyrir að við eigum eftir að framkvæma uppfærslu `i j k`.
- Næst þegar við köllum á `qrec(i, j, ...)` eða `urec(i, j, ...)` þá munum við dreifa uppfærslunni `i j k`.
- Eftir dreifinguna munum við ekki eiga eftir uppfærslu á bilinu $[i, j]$, en við munum eiga eftir uppfærslurnar `i m k` og `(m + 1) j k`.
- Þegar við dreifum uppfærslunni `i i k` þá nægir að uppfæra tilheyrandi lauf í biltrénu.

- Áðan var sagt “laufin geyma þá gildin í listanum og aðrir hnútar geyma summu barna sinna”.
- Þetta gildir ekki fyrir lygn biltré.
- Hnútar lygna biltrjáa þurfa að geyma summu barna sinna, ásamt því að geyma þá summu sem fengist eftir allar óframkvæmdar uppfærslur afkomenda hans.
- Þegar við ferðumst í gegnum tréð til að finna hvert við eigum að setja uppfærsluna uppfærum við tréð jafn óðum.
- Til dæmis, ef við viljum framkvæma uppfærsluna `i j k` þá þurfum við að bæta $k \cdot (j - i + 1)$ við rót biltrésins, því rótin geymir summu allra stakana.

```
void prop(int x, int y, int e)
{
    p[e] += (y - x + 1)*o[e];
    if (x != y) o[LEFT(e)] += o[e], o[RIGHT(e)] += o[e];
    o[e] = 0;
}

int qrec(int i, int j, int x, int y, int e)
{
    prop(i, j, e);
    if (x == i && y == j) return p[e];
    int m = (i + j)/2;
    if (y <= m) return qrec(i, m, x, y, LEFT(e));
    else if (x > m) return qrec(m + 1, j, x, y, RIGHT(e));
    return qrec(i, m, x, m, LEFT(e)) + qrec(m + 1, j, m + 1, y, RIGHT(e));
}
int query(int x, int y)
{
    return qrec(0, p[0] - 1, x, y, 1);
}

void urec(int i, int j, int x, int y, int z, int e)
{
    prop(i, j, e);
    if (x == i && y == j) { o[e] = z; return; }
    int m = (i + j)/2;
    p[e] += (y - x + 1)*z;
    if (y <= m) urec(i, m, x, y, z, LEFT(e));
    else if (x > m) urec(m + 1, j, x, y, z, RIGHT(e));
    else urec(i, m, x, m, z, LEFT(e)), urec(m + 1, j, m + 1, y, z, RIGHT(e));
}
void update(int x, int y, int z)
{
    urec(0, p[0] - 1, x, y, z, 1);
}
```

- Nú hefur `query(...)` sömu tímaflækju og í hefðbundnum biltrjám, það er að segja $\mathcal{O}(\log n)$.
- Loks fæst (með sömu rökum og gefa tímaflækju `query(...)`) að `update(...)` er $\mathcal{O}(\log n)$.

- Til að geta útfært lygnt biltré þurfum við geta skipt uppfærslum í tvennt.
- Við þurfum líka að geta sameinað uppfærslur.
- Í dæminu á undan skiptist `i j k` í `i m k` og `m + 1 j k`.
- Í dæminu á undan sameinast fyrirspurnir `i j x` og `i j y` í `i j x + y`.
- Tökum annað dæmi.
- Við viljum aftur geta reiknað summuna yfir bil.
- Við viljum einnig geta bætt núll við stak x, einum við stak $x + 1$, tveimur við stak $x + 2$, ..., $y - x$ við stak y.
- Þessi uppfærsla svarar til að framkvæma `a[x + i] += i;`, fyrir `i` frá og með núll til og með $y - x$.

- Er hægt að skipta uppfærslunni `x y` í tvennt?
- Nei.
- En við getum bætt við viðfangi sem lýsir hvar summan á að byrja.
- Uppfærslan `x y r` svarar þá til `a[x + i] += r + i;`, fyrir `i` frá og með núll til og með $y - x$.
- Til að framkvæma umbeðna fyrirspurn svörum við þá fyrirspurninni `x y 0`.
- Við getum þá skipt `x y r` í tvennt í `x m r` og `m + 1 y r + m - x + 1`

- Er hægt að sameina uppfærslurnar `x y r` og `x y w`?
- Nei.
- En við getum bætt við öðru viðfangi sem lýsir hversu mikið summan breytist í hverju skrefi.
- Uppfærslan `x y r k` svarar þá til `a[x + i] += r + i*k;`, fyrir `i` frá og með núll til og með $y - x$.
- Til að framkvæma umbeðna fyrirspurn svörum við þá fyrirspurninni `x y 0 1`.
- Við getum nú sameinað tvær fyrirspurnir `x y r k` og `x y z w` í `x y r + z k + w`.
- Við getum einnig skipt `x y r k` í `x m r k` og `m + 1 y r + k*(m - x + 1) k`.

- Til að uppfæra gildin í hverjum hnút nýtum við okkur að uppfærsla `x y r k` bætir í heildina

$$r \cdot (y - x + 1) + k \cdot (y - x + 1) \cdot (y - x)/2$$

við laufin.

```
typedef struct { int v, or, ok; } node;
node p[5*MAXN];

void prop(int x, int y, int e)
{
    p[e].v += (y - x + 1)*p[e].or + p[e].ok*(y - x + 1)*(y - x)/2;
    if (x != y)
        p[RIGHT(e)].or += p[e].or + ((y - x)/2 + 1)*p[e].ok,
            p[LEFT(e)].or += p[e].or, p[LEFT(e)].ok += p[e].ok,
            p[RIGHT(e)].ok += p[e].ok;
    p[e].or = p[e].ok = 0;
}
```

```
int qrec(int i, int j, int x, int y, int e)
{
    prop(i, j, e);
    if (x == i && y == j) return p[e].v;
    int m = (i + j)/2;
    if (y <= m) return qrec(i, m, x, y, LEFT(e));
    else if (x > m) return qrec(m + 1, j, x, y, RIGHT(e));
    return qrec(i, m, x, m, LEFT(e)) + qrec(m + 1, j, m + 1, y, RIGHT(e));
}
int query(int x, int y)
{
    return qrec(0, p[0].v - 1, x, y, 1);
}

void urec(int i, int j, int x, int y, int zr, int e)
{
    prop(i, j, e);
    if (x == i && y == j) { p[e].or = zr; p[e].ok = 1; return; }
    int m = (i + j)/2;
    p[e].v += (y - x + 1)*zr + (y - x + 1)*(y - x)/2;
    if (y <= m) urec(i, m, x, y, zr, LEFT(e));
    else if (x > m) urec(m + 1, j, x, y, zr, RIGHT(e));
    else urec(i, m, x, m, zr, LEFT(e)),
        urec(m + 1, j, m + 1, y, zr + (m - x + 1), RIGHT(e));
}
void update(int x, int y)
{
    urec(0, p[0].v - 1, x, y, 0, 1);
}

void init(int n)
{
    for (int i = 0; i < 5*n; i++) p[i].v = p[i].or = p[i].ok = 0;
    p[0].v = n;
}
```

- Tökum nú dæmi sem er aðeins flóknara.
- Við höfum safn af bilum með heiltölu endapunkta sem byrjar tómt.
- Við getum annað hvort bætt við bili í safnið eða fjarlægt það.
- Við viljum síðan geta fundið lengd sammengis allra bilana í safninu.

- Við getum leyst þetta dæmi með biltréi sem styður aðgerðirnar:
    - Leggja tölu við bil.
    - Finna hversu mörg stök á bilinu eru núll.
- Við munum þurfa að gera ráð fyrir því að það séu aldrei neikvæðar tölur í trénu.
- Þá getum við útfært þetta með biltréi sem geymir minnsta stakið (sem við sáum áðan), ásamt því að geyma hversu oft minnsta stakið kemur fyrir.
- Fjöldi núlla á bili er þá annaðhvort fjöldi minnstu stakanna á bilinu ef minnsta stakið er núll, en núll annars.
- Þetta virkar aðeins ef minnsta stakið er aldrei minna núll.

- Lauf biltrésins munu nú svara til mögulegra gilda endapunkta bilanna.
- Ef endapunktarnir geta verið stórir (til dæmis 10^9) verður þessi lausn of hæg.
- Ef við vitum öll bilin fyrirfram getum við komist hjá því með því að geyma í trénu bara þá endapunkta við notum.
- Gerum ráð fyrir að endapunktar bilanna séu alltaf einhver talnanna $x_1 < x_2 < \cdots < x_m$.
- Til að bæta bilinu $[x_i, x_j]$ í tréð notum við uppfærsluna `i j 1` og til að fjarlægja bilið notum við `i j -1`.
- Gildin í trénu verða aldrei neikvæð því við fjarlægju ekki bil nema að hafa sett það í tréð áður.
- Til að fá rétta lengd geymir laufið sem svarar til vísi i töluna $x_{i+1} - x_i$.
- Við svörum þá fyrirspurninni með `i j - 1` ef minnsta gildið er núll (sem við reiknum með annarskonar fyrirspurn í biltrénu).

```
typedef struct { ll v, o, m; } node;
node p[5*MAXN];

void prop(ll x, ll y, ll e)
{
    p[e].m += p[e].o;
    if (x != y) p[RIGHT(e)].o += p[e].o, p[LEFT(e)].o += p[e].o;
    p[e].o = 0;
}

ll qrecm(ll i, ll j, ll x, ll y, ll e)
{
    prop(i, j, e);
    if (x == i && y == j) return p[e].m;
    ll m = (i + j)/2;
    if (y <= m) return qrecm(i, m, x, y, LEFT(e));
    else if (x > m) return qrecm(m + 1, j, x, y, RIGHT(e));
    return min(qrecm(i, m, x, m, LEFT(e)),
            qrecm(m + 1, j, m + 1, y, RIGHT(e)));
}
ll querymin(ll x, ll y)
{
    return qrecm(0, p[0].v - 1, x, y, 1);
}
```

```
ll qrec(ll i, ll j, ll x, ll y, ll e)
{
    prop(i, j, e);
    if (x == i && y == j) return p[e].v;
    ll m = (i + j)/2, v1, v2, w1, w2;
    if (y <= m) return qrec(i, m, x, y, LEFT(e));
    else if (x > m) return qrec(m + 1, j, x, y, RIGHT(e));
    v1 = qrec(i, m, x, m, LEFT(e)), v2 = qrec(m + 1, j, m + 1, y, RIGHT(e));
    w1 = qrecm(i, m, x, m, LEFT(e)), w2 = qrecm(m + 1, j, m + 1, y, RIGHT(e));
    return (w1 <= w2)*v1 + (w1 >= w2)*v2;
}
ll query(ll x, ll y)
{
    if (y < x) return 0;
    ll z = qrec(0, p[0].v - 1, x, y, 1);
    return querymin(x, y) == 0 ? z : 0;
}

void urec(ll i, ll j, ll x, ll y, ll z, ll e)
{
    prop(i, j, e);
    if (x == i && y == j) { p[e].o = z; return; }
    ll m = (i + j)/2, v1, v2;
    if (y <= m) urec(i, m, x, y, z, LEFT(e));
    else if (x > m) urec(m + 1, j, x, y, z, RIGHT(e));
    else urec(i, m, x, m, z, LEFT(e)), urec(m + 1, j, m + 1, y, z, RIGHT(e));
    v1 = p[LEFT(e)].m + p[LEFT(e)].o, v2 = p[RIGHT(e)].m + p[RIGHT(e)].o;
    p[e].m = min(v1, v2);
    p[e].v = (v1 <= v2)*p[LEFT(e)].v + (v1 >= v2)*p[RIGHT(e)].v;
}
void update(ll x, ll y, ll z)
{
    urec(0, p[0].v - 1, x, y, z, 1);
}
```

```
void irec(ll i, ll j, ll e, ll *a)
{
    p[e].v = p[e].o = p[e].m = 0;
    if (i == j) { p[e].v = a[i + 1] - a[i]; return; }
    ll m = (i + j)/2;
    irec(i, m, LEFT(e), a), irec(m + 1, j, RIGHT(e), a);
    p[e].v = p[LEFT(e)].v + p[RIGHT(e)].v;
}
void init(ll *a, ll n)
{
    irec(0, (p[0].v = n) - 1, 1, a);
}
```