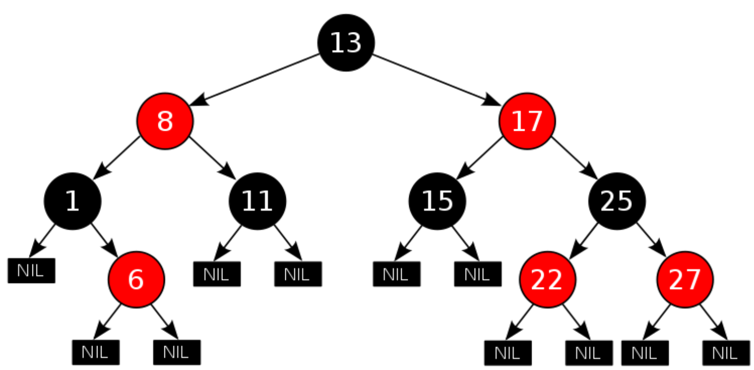
Crveno-crna drveta

Crveno-crno drvo (eng. red-black tree, RBT) je uređeno binarno drvo koje dodatno zadovoljava naredne invarijante:

1. Svaki čvor je ili crven ili crn.
2. Koren je crn.
3. Svi listovi su crni i ne sadrže vrednosti (označavamo ih sa NIL).
4. Svi crveni čvorovi imaju tačno dva crna deteta.
5. Sve putanje od nekog čvora do listova u njegovom poddrvetu sadrže isti broj crnih čvorova.

Crveno-crna drveta je 1972. godine izumeo Rudolf Bayer, dok su se nešto kasnije Leonidas Guibas i Robert Sedgewick bavili analizom njihovih svojstava i uveli crveno-crnu konvenciju bojenja čvorova[1]. 1993. godine Arne Andersson je naknadno osmislio jednostavniju varijantu crveno-crnih drveta kojima se pojednostavljuju operacije umetanja i brisanja.

Primer jednog crveno-crnog drveta je prikazan na slici 1. Primetimo da je koren drveta crne boje, a da su svi listovi crne boje i sadrže vrednost NIL. Svaki crveni čvor ima tačno dva crna deteta, dok za crne čvorove važi da mogu imati dva crvena, dva crna, ili crveno i crno dete. Razmotrimo sve putanje od čvora sa vrednošću 17 do listova – sve putanje sadrže po 2 crna čvora, ali potencijalno različit broj crvenih čvorova.



Slika 1: Primer crveno-crnog drveta.

NIL vrednosti ne moraju nužno biti kodirane kao zasebni čvorovi drveta, već bi mogle da odgovaraju vrednostima pokazivača. Međutim, ovakav pogled na njih olakšava izvođenje operacija sa crveno-crnim drvetom.

[1] Uz crnu, crvena boja je izabrana zato što je davala najbolji prikaz na laserskom štampaču koji je autorima bio dostupan.

Prethodno navedena svojstva nam garantuju da će *putanja od korena do njemu najudaljenijeg lista biti najviše duplo duža nego putanja do njemu najbližeg lista.* Zaista, najkraća putanja do nekog lista će se sastojati samo od crnih čvorova, dok će se se u najdužoj putanji naizmenično smenjivati crveni i crni čvorovi (jer na osnovu uslova 4. posle crvenog čvora mora doći crni, a na osnovu uslova 5. sve putanje do listova imaju isti broj crnih čvorova). Ovo svojstvo nam garantuje određeni vid balansiranosti crveno-crnog drveta.

Dokažimo da su prethodno navedeni uslovi dovoljni da bi se garantovalo da visina crveno-crnog drveta logaritamski zavisi od broja čvorova. Odatle bi sledila logaritamska složenost osnovnih operacija (pod uslovom da svako umetanje u drvo i brisanje elementa iz drveta održava nabrojanih pet invarijanti).

Neka je:

- $h(v)$ – *visina* poddrveta čiji je koren čvor v, tj. broj čvorova od čvora v do najudaljenijeg lista (ne računajući čvor v);
- $h_b(v)$ – *crna visina* poddrveta čiji je koren čvor v, tj. broj crnih čvorova od čvora v do proizvoljnog lista (ne računajući čvor v ako je on crn)[2].

Na primer, visina h čvora sa vrednošću 13 sa slike 1 je 4, a crna visina h_b je 2. Slično, visina čvora sa vrednošću 8 jednaka je 3, a crna visina 2.

Pod *unutrašnjim čvorovima* drveta podrazumevaćemo sve čvorove sem listova, odnosno NIL čvorova. Mi ćemo se u daljem tekstu uglavnom fokusirati na unutrašnje čvorove drveta jer oni sadrže vrednosti ključeva.

Naredna lema nam daje vezu između broja unutrašnjih čvorova i crne visine crveno-crnog drveta.

Lema: Crveno-crno drvo sa korenom u čvoru v ima bar $2^{h_b(v)} - 1$ unutrašnjih čvorova.

Dokaz: Označimo sa n broj unutrašnjih čvorova datog drveta. Dokaz teče indukcijom po visini drveta tj. po vrednosti $h(v)$.

- Bazu čini slučaj $h(v) = 0$. Visinu nula ima jedino NIL čvor za koji važi $n = 0$, pa je tada i $h_b(v) = 0$, odnosno $2^{h_b(v)} - 1 = 0$, pa je zahtev tvrđenja trivijalno ispunjen.

- Pretpostavimo kao induktivnu hipotezu da svako drvo sa korenom u čvoru w čija je visina $h(w) < k$ ima bar $2^{h_b(w)} - 1$ unutrašnjih čvorova. Neka drvo ima koren u čvoru v i neka je $h(v) = k$. Pošto je $h(v) > 0$ čvor v je unutrašnji i ima dva deteta v_l i v_d koji su koreni drveta visine manje od k te za njih važi induktivna hipoteza. Za čvor v_l važi da mu je crna visina ili $h_b(v)$ (ako je čvor v_l crven) ili $h_b(v) - 1$ (ako je čvor v_l crn). Analogno važi i za čvor v_d. Na osnovu induktivne hipoteze za broj unutrašnjih čvorova

[2] Prema svojstvu 5. pojam crne visine je dobro definisan jer sve putanje od nekog čvora do listova sadrže isti broj crnih čvorova.

n_l poddrveta sa korenom u čvoru v_l važi:

$$n_l \geq 2^{h_b(v_l)} - 1 \geq 2^{h_b(v)-1} - 1,$$

i slično za broj unutrašnjih čvorova n_d poddrveta sa korenom u čvoru v_d:

$$n_d \geq 2^{h_b(v)-1} - 1,$$

pa je broj unutrašnjih čvorova drveta sa korenom u čvoru v bar

$$n = n_l + n_d + 1 = 2(2^{h_b(v)-1} - 1) + 1 = 2^{h_b(v)} - 1$$

(objedinili smo unutrašnje čvorove u oba poddrveta i čvor v). $\square$

Naredna teorema nam daje vezu između broja unutrašnjih čvorova crveno-crnog drveta i njegove visine.

Teorema: Visina h crveno-crnog drveta koje sadrži n unutrašnjih čvorova zadovoljava uslov $h \leq 2\log_2(n+1)$.

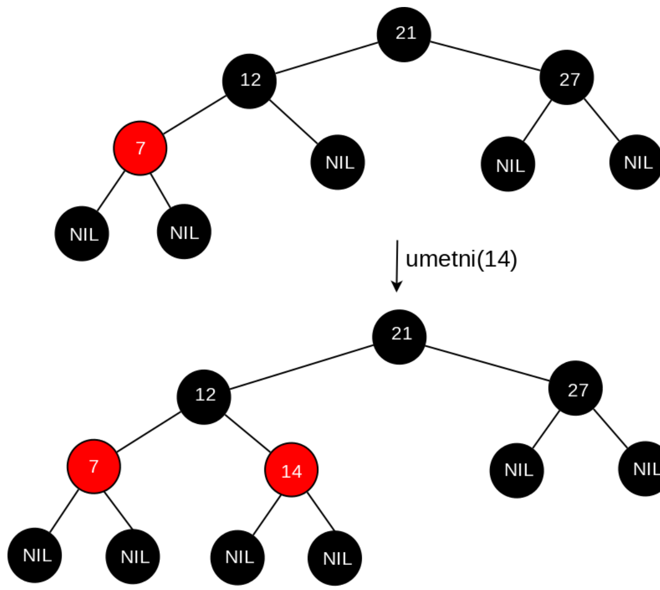
Dokaz: Na osnovu prethodne leme važi da je broj unutrašnjih čvorova u crveno-crnom drvetu sa korenom u čvoru v bar $2^{h_b(v)} - 1$. Pošto je bar pola čvorova na svakoj putanji od korena v do listova crno, važi da je $h_b(v) \geq h(v)/2$. Zato je na osnovu prethodne leme broj unutrašnjih čvorova n veći ili jednak $2^{h(v)/2} - 1$. Odavde sledi $n + 1 \geq 2^{h(v)/2}$, pa je $\log_2(n+1) \geq h(v)/2$ i važi da je $h(v) \leq 2\log_2(n+1)$. $\square$

Direktna posledica tvrđenja prethodne teoreme je da je visina crveno-crnog drveta koje sadrži n čvorova $O(\log n)$. Odavde sledi da će operacija pretrage u crveno-crnom drvetu biti logaritamske složenosti u funkciji broja elemenata u drvetu.

U nastavku teksta razmotrićemo na koji način se u crveno-crnim drvetima realizuju operacije koje menjaju njegovu strukturu, poput umetanja i brisanja elemenata. Poseban akcenat je kao i kod AVL drveta stavljen na to da se nakon umetanja, odnosno brisanja elementa iz crveno-crnog drveta mogu pokvariti neka od svojstava crveno-crnog drveta i tada je neophodno izmeniti boje čvorova i/ili preusmeriti neke od pokazivača, tako da dobijeno drvo i dalje ostane crveno-crno. Preusmeravanje pokazivača se vrši operacijama rotacije drveta. Postoje *leva* i *desna rotacija* i one odgovaraju LL i RR rotaciji u AVL drvetu.

Umetanje elementa u crveno-crno drvo Umetanje novog čvora u crveno-crno drvo se sprovodi na način uobičajen za uređena binarna drveta: na osnovu vrednosti ključa pronalazi se mesto gde treba dodati novi čvor i na mesto NIL lista se postavlja novi čvor sa zadatom vrednošću ključa. Pritom se novi čvor boji crveno i dodaju mu se dva crna NIL deteta. Na slici 2 ilustrovan je primer umetanja elementa u crveno-crno drvo. Primetimo da je drvo i nakon umetanja elementa ostalo crveno-crno.

Razmotrimo koja svojstva crveno-crnih drveta su pri umetanju elementa i njegovog bojenja u crveno mogla biti narušena. Svojstvo 1, kojim se tvrdi da su svi



Slika 2: Primer umetanja elementa sa vrednošću 14 u crveno-crno drvo kojim se ne narušava nijedan od uslova.

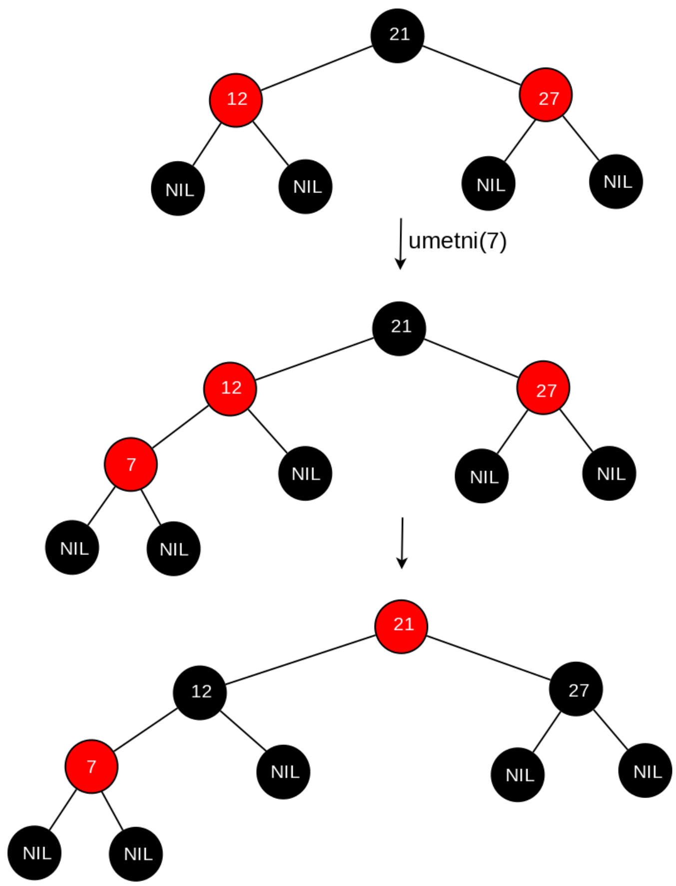
čvorovi obojeni crveno ili crno, i svojstvo 3 kojim se tvrdi da je svaki list crn trivijalno ostaju očuvana. Svojstvo 5, kojim se tvrdi da je broj crnih čvorova isti na svakoj putanji od fiksiranog čvora do proizvoljnog lista je zadovoljen jer je novi čvor obojen crveno i sa svoja dva crna deteta (lista) je zamenio prethodni crni list. Stoga, jedina dva svojstva koja su mogla biti narušena operacijom umetanja elementa su svojstva 2 i 4. Svojstvo 2, kojim se zahteva da je koren drveta crn je mogao biti narušen samo u slučaju da je novododati čvor koren drveta, dok je svojstvo 4 kojim se zahteva da crveni čvor ne može imati crveno dete mogao biti narušen ako je roditelj novododatog čvora bio crven.

Važno je primetiti da je tačno jedno od svojstava 2 i 4 moglo biti narušeno: naime, svojstvo 2 je moglo biti narušeno samo u situaciji kada je čvor koji se umeće koren drveta, ali on u ovom slučaju nema roditelja, a ima dva crna NIL deteta, te svojstvo 4 nije moglo biti istovremeno narušeno.

U daljem razmatranju ćemo često referisati na čvor koji je brat roditelja razmatranog čvora i njega ćemo zvati *ujakom* datog čvora.

Razmotrimo kako održati sva svojstva crveno-crnih drveta nakon umetanja novog elementa z. Najpre, ukoliko je narušen uslov 2, odnosno ako se čvor umeće u prazno drvo i postaje koren drveta, potrebno je čvoru koji se umeće boju promeniti iz crvene u crnu.

Inače ćemo pretpostaviti da je roditelj čvora z levo dete svog oca (slučaj kada je roditelj čvora z desno dete svog oca se analogno razmatra). Odnos koji je narušen nakon umetanja jeste odnos novog čvora z i njegovog roditelja, jer su oba ova čvora crvena. Razlikujemo tri različita slučaja:



Slika 3: Primer umetanja elementa u crveno-crno drvo kada su i roditelj i ujak novog čvora crveni.

1. Prvi slučaj nastupa kada su istovremeno i roditelj i ujak čvora z crvene boje. S obzirom na to da je deda čvora z nužno obojen crno, možemo njegovoj deci (roditelju i ujaku čvora z) promeniti boju iz crvene u crnu, a njega obojiti u crveno (slika 3). Na ovaj način smo razrešili problem jer su istovremeno i čvor z i njegov roditelj bili crvene boje, a održali smo i svojstvo 5: broj crnih čvorova na putanjama ostaje nepromenjen. Pošto je nakon ove popravke boja dede čvora z iz crne izmenjena u crvenu, mogli smo na ovaj način poremetiti svojstvo 4 između dede čvora z i njegovog roditelja, te isti postupak treba sada sprovesti za čvor dva nivoa iznad čvora z, odnosno za dedu čvora z. Međutim, s obzirom na to da je visina drveta logaritamske složenosti u funkciji broja čvorova drveta, a da se svaka

od popravki može izvesti u složenosti $O(1)$, ukupna složenost popravke nakon umetanja je odozgo ograničena sa $O(\log n)$.

2. Drugi mogući slučaj se odnosi na situaciju kada je ujak čvora z crn, a z je desno dete svoga roditelja (slika 4(a)). U ovom scenariju se primenom leve rotacije poddrvo čiji je koren roditelj čvora z transformiše u drvo u kome je z levo dete svog roditelja i na taj način se svodi na treći slučaj. S obzirom na to da su i čvor z i njegov roditelj crvene boje, rotacija ne utiče na crne visine čvorova, te stoga ni na svojstvo 5.

3. Treći slučaj nastupa kada je ujak čvora z crn, a z je levo dete svoga roditelja. Primetimo da je i u ovom slučaju deda čvora z nužno obojen crno. U ovom slučaju vrši se izmena boja čvorova dede i roditelja čvora z – deda postaje crven, a roditelj crn. Nakon toga vrši se desna rotacija poddrveta sa korenom u roditelju čvora z. Na ovaj način se čuva svojstvo 5 (slika 4(b)). Primetimo da nakon ovoga ne postoje više dva crvena čvora koja su u odnosu roditelj-dete, niti se neki od problema potencijalno prolongira uz drvo. Dakle, u ovom slučaju je jedna rotacija ovog tipa bila dovoljna da drvo ponovo zadovoljava svih pet navedenih svojstava.

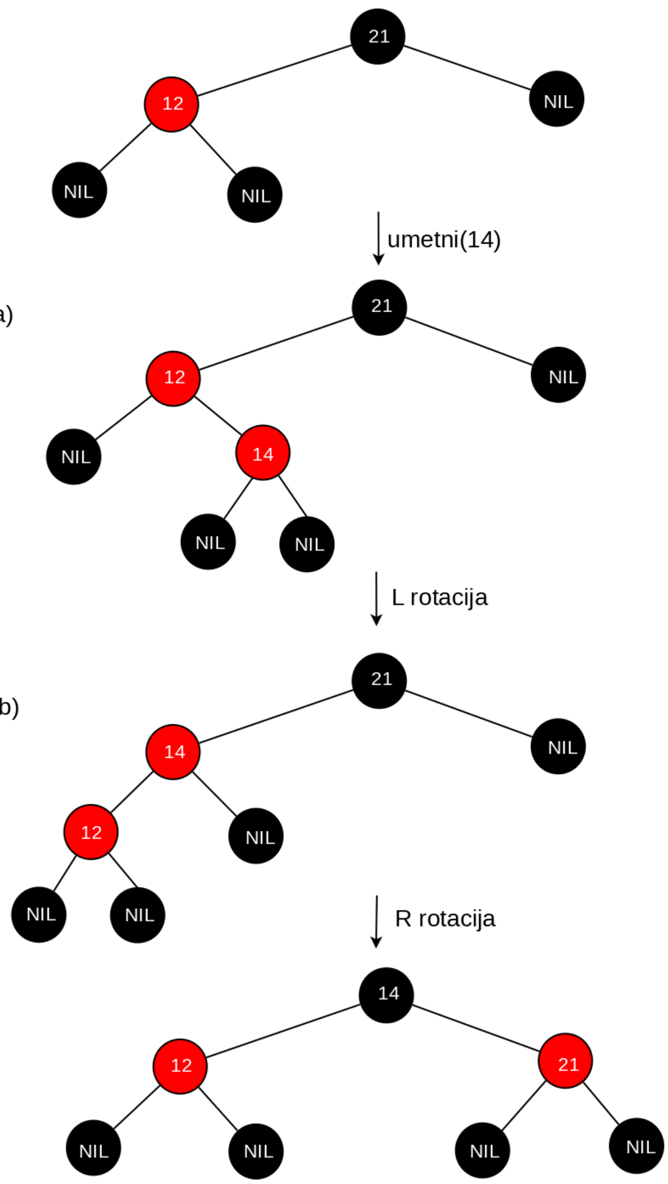
Primetimo da nijedan od tri razmatrana scenarija ne narušava svojstva $1, 3$ i 5, a ni svojstvo 2 jer čvor z nije koren. Primetimo takođe, da se samo u slučaju prvog od tri različita slučaja umetanja čvora, čvor z izdiže dva nivoa iznad i ponovo se razmatra da li je došlo do narušavanja nekog svojstva crveno-crnih drveta. S obzirom na to da je visina crveno-crnog drveta odozgo ograničena sa $O(\log n)$, može se izvesti najviše $O(\log n)$ korekcija, pa je ukupna složenost popravki u najgorem slučaju $O(\log n)$.

Brisanje elementa iz crveno-crnog drveta Kao i u slučaju AVL drveta, operacija brisanja je složenija i nećemo je detaljnije razmatrati u ovom materijalu.

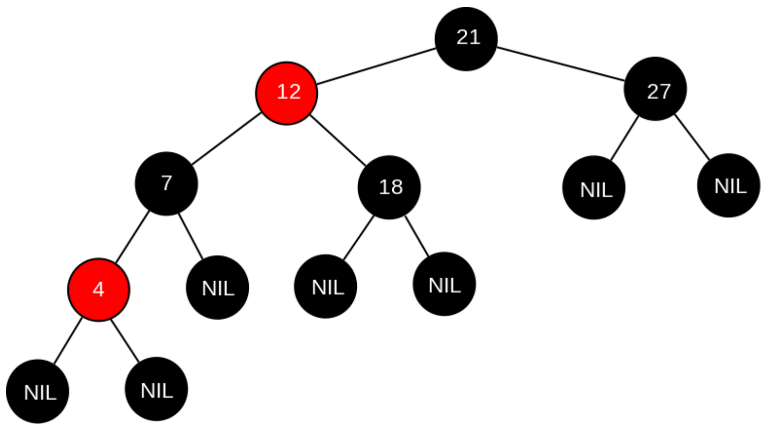
Poređenje AVL drveta sa crveno-crnim drvetima AVL drveta su strožije izbalansirana od crveno-crnih drveta. U opštem slučaju visina AVL drveta je manja, pa se pretraga u AVL drvetu izvodi efikasnije nego u crveno-crnim drvetima. Stoga se AVL drveta pokazuju kao bolja opcija kada je očekivani broj operacija pretraživanja značajno veći od broja operacija umetanja. Međutim, bojenje čvorova u crveno-crnim drvetima omogućava manji broj potrebnih operacija rebalansiranja, jer nam bojenje čvorova nekada omogućava da izbegnemo ili makar smanjimo broj operacija rebalansiranja. AVL drveta sobom nose veće troškove rotacije nego crveno-crna drveta, što će, u situacijama kada se očekuje veliki broj umetanja, dovesti do velikog broja rotacija i tada neka druga struktura podataka poput crveno-crnih drveta predstavlja bolji izbor.

Primetimo i to da se svako AVL drvo može transformisati u crveno-crno bojenjem čvorova redom u crveno i crno, dok postoje crveno-crna drveta koja nisu AVL (slika 5).

Opišimo način na koji se proizvoljno AVL drvo može obojiti tako da bude

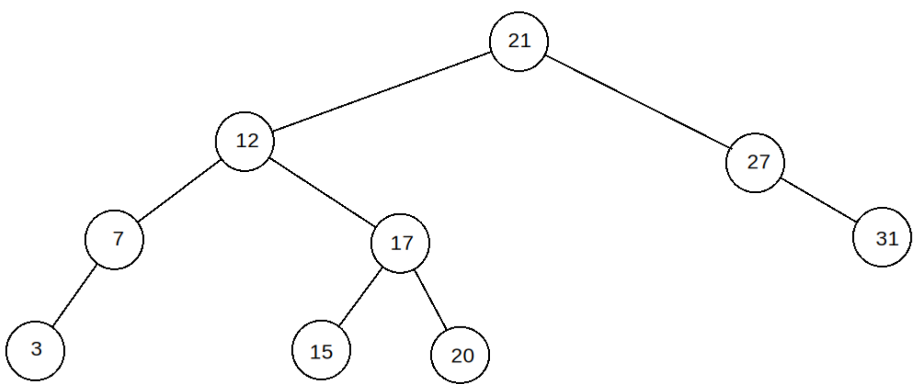


Slika 4: Primer umetanja elementa u crveno-crno drvo kada je roditelj novog čvora crven, a ujak novog čvora crn.



Slika 5: Primer crveno-crnog drveta koje nije AVL.

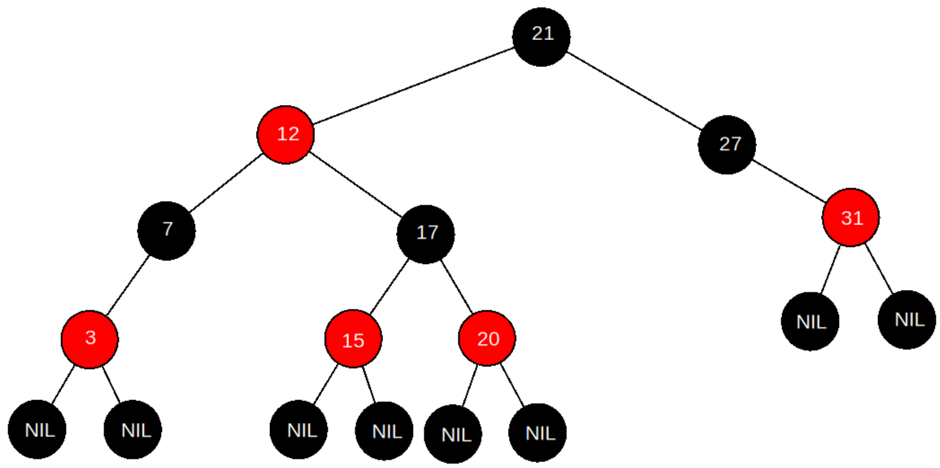
crveno-crno drvo. Razmotrimo primer AVL drveta sa slike 6. U duhu definicije crveno-crnog drveta, razmatraćemo da svaki od listova AVL drveta ima dva deteta u crveno-crnom drvetu, koji sadrže NIL vrednosti.



Slika 6: AVL drvo.

Ključno svojstvo koje je potrebno obezbediti jeste da svaka putanja od proizvoljnog čvora do lista ima istu crnu dubinu. Ono što omogućava da se ovako nešto postigne jeste invarijanta AVL drveta po kojoj se visine levog i desnog podstabla u odnosu na proizvoljni čvor razlikuju najviše za jedan.

Koren drveta inicijalno bojimo crnom bojom i nakon toga pozivamo funkciju bojenja za njegovu decu. Ako je visina levog drveta manja od visine desnog ili je visina levog drveta parna, levog sina bojimo crno i rekurzivno pozivamo bojenje njegove dece. Inače, levog sina bojimo crveno i rekurzivno pozivamo bojenje njegove dece. Nakon toga razmatramo desno dete i, slično, ako je visina desnog deteta manja od visine levog ili je visina desnog deteta parna, desno dete bojimo crno; inače, desno dete bojimo crveno. AVL drvo sa slike 6 obojeno na ovaj način prikazano je na slici 7.



Slika 7: Crveno-crno drvo.

Na ovaj način drvo obilazimo u dubinu i bojimo svaki od čvorova drveta, te je svojstvo 1 zadovoljeno. Takođe, bojenjem korena u crno zadovoljeno je i svojstvo 2. Primetimo da ovakvim bojenjem su crveni čvorovi uvek neparne visine, a pošto su listovi visine 0, biće obojeni crno te važi i uslov 3. Deca crvenih čvorova će biti ili parne visine ili će biti manje visine od bratskog čvora, tako da će biti obojena crno, te važi i uslov 4. Svojstvo 5 se može dokazati indukcijom po visini drveta.

Zadaci za vežbu

1. Prikazati crveno-crno drvo koje se dobija uzastopnim umetanjima ključeva $41, 38, 31, 12, 19, 8$ u inicijalno prazno crveno-crno drvo.

2. Nacrtati potpuno uređeno binarno drvo visine 3 koje sadrži vrednosti ključeva $\{1, 2, \ldots, 15\}$. Dodati NIL čvorove i obojiti čvorove na tri različita načina tako da je crna visina drveta redom 2, 3 i 4.

3. Koliki je maksimalni, a koliki minimalni mogući broj unutrašnjih čvorova u crveno-crnom drvetu čija je crna visina k?

4. Opisati crveno-crno drvo sa n ključeva koji ima maksimalni mogući odnos crvenih unutrašnjih čvorova u odnosu na crne unutrašnje čvorove. Koliki je taj odnos?

5. Opisati crveno-crno drvo sa n ključeva koji ima minimalni mogući odnos crvenih unutrašnjih čvorova u odnosu na crne unutrašnje čvorove. Koliki je taj odnos?

6. Objasniti zašto u crveno-crnom drvetu crveni čvor ne može imati tačno jedno dete koji nije NIL čvor.
7. Ukoliko bi se čvor čije se umetanje vrši bojio u crno umesto u crveno, onda se svojstvo 4 ne bi moglo narušiti. Zašto ne radimo na ovaj način?
8. Da li crveno-crno drvo može da sadrži samo crne čvorove? Ukoliko ne može, obrazložiti zašto ne može, a ako može precizirati kako izgleda.