# Strukture podataka

## Balansirana uređena binarna drveta

Za neko binarno drvo kažemo da je *uređeno* ako je prazno, ili ako su njegovo levo i desno poddrvo uređeni i vrednost u korenu je veća ili jednaka od vrednosti svih čvorova u levom poddrvetu, a manja od vrednosti svih čvorova u desnom poddrvetu. Drugim rečima, za svaki unutrašnji čvor u drvetu treba da važi da je njegova vrednost veća ili jednaka od svih vrednosti u levom, a manja od svih vrednosti u desnom poddrvetu. Istaknimo da uslov da je vrednost u svakom čvoru veća ili jednaka od vrednosti levog deteta, a manja od vrednosti desnog deteta nije dovoljan da bi drvo bilo uređeno.

Uređena binarna drveta su struktura podataka pogodna za implementaciju asocijativnih struktura podataka poput uređenih mapa i uređenih skupova[1]. Međutim, treba imati na umu da su ona efikasna samo kada su *balansirana*. Naime, ukoliko uređeno binarno drvo nije balansirano, operacije pretrage, umetanja i brisanja elemenata mogu zahtevati linearno vreme u odnosu na broj elemenata u drvetu. Na primer, ako u prazno drvo vršimo redom umetanje $n$ elemenata čije su vrednosti ključeva strogo rastuće ili strogo opadajuće, dobija se uređeno binarno drvo visine $n$. Postavlja se pitanje da li je moguće dizajnirati uređeno binarno drvo koje bi imalo garantovanu visinu $O(\log n)$, bez obzira na redosled umetanja i brisanja elemenata. Ako bi tako nešto bilo moguće, onda bi vremenska složenost operacija pretrage, umetanja i brisanja elementa u najgorem slučaju bila logaritamska u odnosu na broj elemenata u drvetu.

Kako bi se drvo održavalo balasiranim prilikom izvršavanja operacija koje menjaju strukturu drveta, poput umetanja ili brisanja elementa, potrebno je sprovesti dodatne operacije *balansiranja drveta*. Pritom, moramo biti veoma pažljivi da prilikom balansiranja drveta ne narušimo uslov uređenosti drveta.

Zbog velikog značaja koji imaju, razvijene su različite varijante uređenih binarnih drveta koje koriste različite algoritme balansiranja. One se međusobno razlikuju i po načinu na koji definišu balansiranost, tj. invarijantama koje održavaju, kao i u tome kada i kako vrše balansiranje.

U nastavku ćemo razmotriti dve najčešće korišćene vrste balansiranih uređenih binarnih drveta:

- Adeljson-Veljski Landisova drveta (AVL drveta) i
- crveno-crna drveta.

Pored navedenih, postoje i druge vrste balansiranih uređenih drveta, kao što su 2-3 drveta, 2-3-4 drveta, B drveta, B+ drveta i druga.

[1] Asocijativne strukture podataka podrazumevaju pristup elementima ne na osnovu indeksa, odnosno pozicije elementa u strukturi, već na osnovu vrednosti ključa.

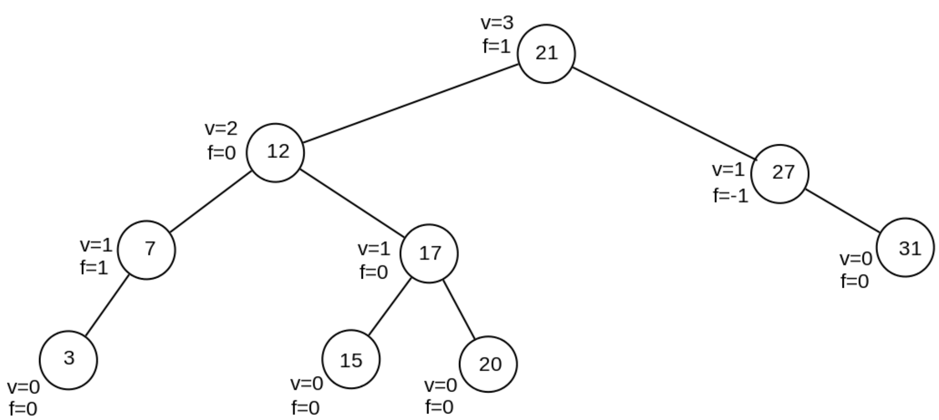
**AVL drveta**

*AVL drvo* predstavlja jednu vrstu balansiranih uređenih binarnih drveta. AVL drvo je dobilo naziv po prvim slovima prezimena njegovih autora: Georgii Adelson-Velskii i Evgenii Mikhailovich Landis. Za AVL drvo važi da složenost ni jedne od operacija traženja, umetanja i brisanja u najgorem slučaju nije veća od $O(\log n)$, gde je $n$ broj elemenata u drvetu. To se postiže tako što se posle svake od operacija umetanja i brisanja elemenata ulaže dodatni napor da se drvo izbalansira, tako da visina drveta uvek bude $O(\log n)$. Pri tome se balansiranost drveta definiše tako da se može lako održavati. Neformalno, ideja AVL drveta je da za svaki čvor važi da se su njegovo levo i desno približnih visina.

*Visina drveta* je najveće rastojanje korena do nekog lista. Ona se može razmatrati u terminima broja čvorova ili broja grana na najdužem putu od korena do lista. Mi ćemo ovde usvojiti drugu konvenciju, odnosno pod visinom drveta podrazumevati maksimalni broj grana na putu od korena do lista. Odatle sledi da drvo koje sadrži jedan čvor ima visinu 0, drvo koje sadrži dva čvora ima visinu 1, dok je prazno drvo visine $-1$.

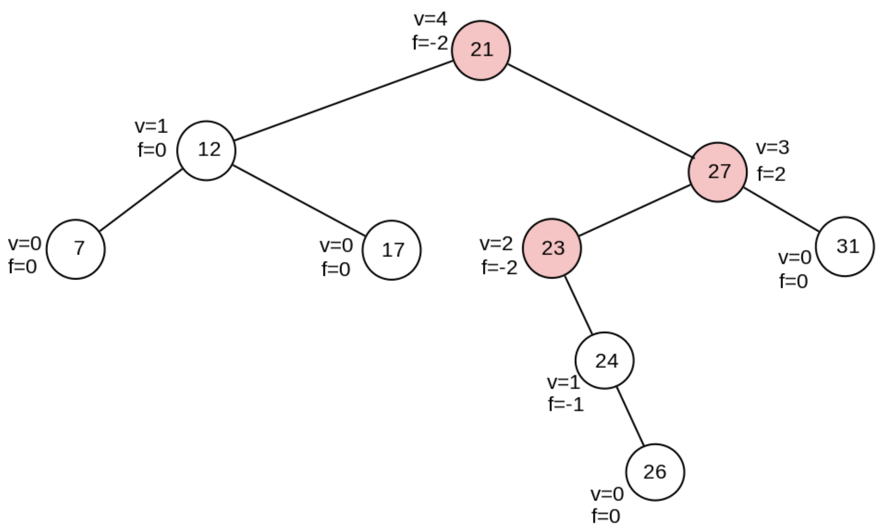
*Faktor ravnoteže čvora* je razlika visina levog i desnog poddrveta tog čvora. Pritom ako čvor nema levo (desno) dete, visina levog (desnog) poddrveta je jednaka $-1$, kao visina praznog drveta.

AVL drvo se definiše kao uređeno binarno drvo kod koga je *za svaki čvor apsolutna vrednost faktora ravnoteže manja ili jednaka od jedan*; dakle za svaki čvor važi da se visine levog i desnog poddrveta mogu razlikovati najviše za jedan.




Slika 1: Primer AVL drveta. Uz svaki čvor drveta prikazana je njegova visina $v$ i faktor ravnoteže $f$ tog čvora.

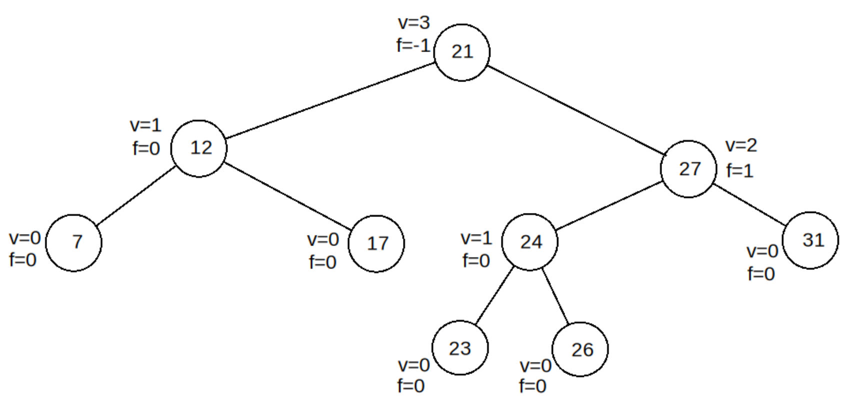
Na slici 1 dat je primer jednog AVL drveta. Uz svaki čvor prikazana je visina drveta sa korenom u tom čvoru i faktor ravnoteže čvora. Možemo primetiti da su faktori ravnoteže svih čvorova datog drveta iz skupa $\{-1, 0, 1\}$. Na primer, faktor ravnoteže čvora čija je vrednost ključa 7 je 1 jer je visina njegovog levog




Slika 2: Primer uređenog binarnog drveta koje nije AVL drvo.

poddrveta 0, a desnog $-1$, dok je faktor ravnoteže čvora sa vrednošću ključa 21 jednak 1, jer je visina njegovog levog poddrveta 2, a desnog 1.

Na slici 2 prikazano je jedno uređeno binarno drvo koje nije AVL drvo. Naime, faktori ravnoteže čvorova sa vrednostima ključa redom 21, 27 i 23 nisu iz skupa $\{-1, 0, 1\}$. Međutim, malim korekcijama ovog drveta možemo od njega dobiti AVL drvo prikazano na slici 3.




Slika 3: Transformacija prethodno ilustrovanog ne-AVL drveta u AVL drvo operacijom balansiranja drveta.

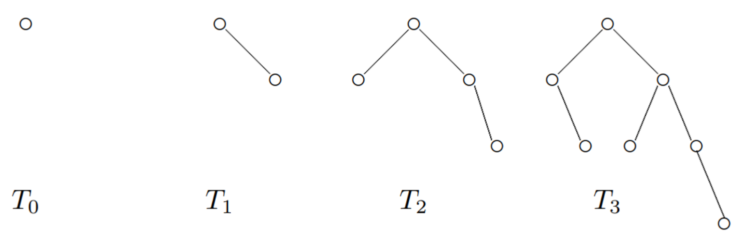
Prema definiciji AVL drveta uslov za faktor ravnoteže treba da važi za svaki čvor u drvetu, odakle sledi da je svako poddrvo AVL drveta takođe AVL drvo.

Razmotrimo da li je uslov da je apsolutna vrednost faktora ravnoteže svakog čvora u drvetu manja ili jednaka 1 dovoljno jak da garantuje da je visina AVL drveta od $n$ elemenata $O(\log n)$.

**Teorema:** Visina $h$ AVL drveta sa $n$ čvorova zadovoljava uslov $h \leq 2\log_2 n$.

**Dokaz:** Dokažimo ovo tvrđenje indukcijom po visini drveta $h$. Prvo, primetimo da visina drveta ne određuje jednoznačno broj čvorova u drvetu. Za dokazivanje navedene nejednakosti kritična su drveta sa *najmanjim* brojem čvorova za datu visinu (jer ako ona ne krše nejednakost, ne krše je ni drveta sa više čvorova zbog monotonosti logaritamske funkcije). Stoga se u dokazu bavimo najmanjim drvetima date visine.

Neka je $T_h$ AVL drvo visine $h \geq 0$ sa najmanjim mogućim brojem čvorova. Drvo $T_0$ sadrži samo koren, a drvo $T_1$ koren i jedno dete, recimo desno. Za drvo $T_h$ gde je $h \geq 2$ važi da se visina levog i desnog deteta mogu razlikovati najviše za 1, te jedno dete (npr. desno) mora biti visine $h-1$, dok je drugo (npr. levo) visine $h-2$. Pritom, da bi dato AVL drvo imalo najmanji mogući broj čvorova, onda i njegovo poddrvo manje visine $h-2$ takođe treba da bude AVL drvo sa minimalnim brojem čvorova, dakle $T_{h-2}$; slično, njegovo drugo poddrvo treba da bude $T_{h-1}$. Drveta opisana ovakvom rekurentnom jednačinom nazivaju se *Fibonačijeva drveta*. Nekoliko prvih Fibonačijevih drveta, takvih da je u svakom unutrašnjem čvoru visina levog poddrveta manja, prikazano je na slici 4.



Slika 4: Fibonačijeva AVL drveta.

Označimo sa $n_h$ minimalni broj čvorova AVL drveta visine $h$, tj. broj čvorova drveta $T_h$. Važi $n_0 = 1$, $n_1 = 2$, $n_2 = 4, \ldots$ Prema definiciji Fibonačijevih drveta važi $n_h = n_{h-1} + n_{h-2} + 1$, za $h \geq 2$. S obzirom na to da je $n_h > n_{h-1}$ za svako $h \geq 1$ važi:

$$n_h = n_{h-1} + n_{h-2} + 1 > 2n_{h-2} + 1 > 2n_{h-2} \tag{1}$$

Dokažimo indukcijom po visini drveta $h$ da važi tvrđenje:

$$n_h \geq 2^{h/2}$$

Za $h = 0$ i $h = 1$ tvrđenje trivijalno važi. Pretpostavimo da tvrđenje važi za $h-2$, odnosno da važi $n_{h-2} \geq 2^{(h-2)/2}$ i dokažimo da onda važi i za $h$. Na osnovu veze (1) važi

$$n_h \geq 2 \cdot 2^{(h-2)/2} = 2^{h/2-1+1} = 2^{h/2}$$

te tvrđenje važi. Nakon logaritmovanja obe strane nejednakosti dobijamo

$$h \leq 2 \log_2 n_h \leq 2 \log_2 n$$

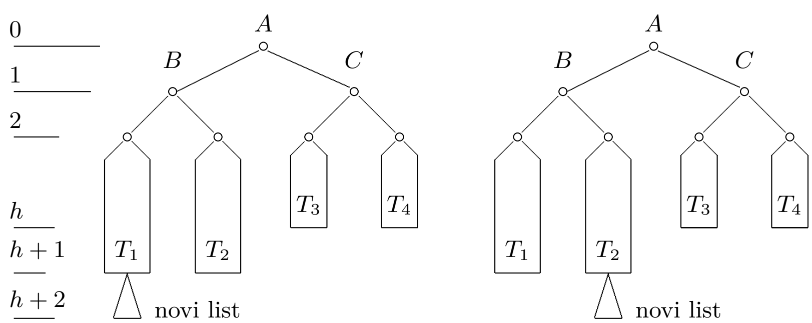
jer je, po pretpostavci, za sva AVL drveta visine $h$ broj čvorova $n$ veći ili jednak od $n_h$.

Dakle, visina AVL drveta sa $n$ čvorova je $O(\log n)$.$\square$

Pošto je visina AVL drveta od $n$ elemenata jednaka $O(\log n)$, operacija pretrage biće logaritamske složenosti. Ono što preostaje da se pokaže jeste na koji se način se realizuju operacije umetanja i brisanja elementa iz AVL drveta i kako se nakon svake operacije umetanja i brisanja održava uslov balansiranosti drveta.

**Umetanje elementa u AVL drvo** Prilikom umetanja novog elementa u AVL drvo postupa se najpre na način uobičajen za uređeno binarno drvo: pronalazi se mesto čvoru, pa se u drvo dodaje novi list sa ključem jednakim zadatoj vrednosti. Čvorovima na putu koji se tom prilikom prelaze odgovaraju faktori ravnoteže iz skupa $\{0, 1, -1\}$. Posebno je interesantan poslednji čvor na tom putu koji ima faktor ravnoteže različit od nule, tzv. *kritični čvor*. Ispostavlja se da je prilikom umetanja elementa u AVL drvo *dovoljno izbalansirati poddrvo sa korenom u kritičnom čvoru*. To se postiže različitim tipovima *rotacija drveta*, koje vrše reorganizaciju čvorova u drvetu, održavajući pritom uređenost drveta.

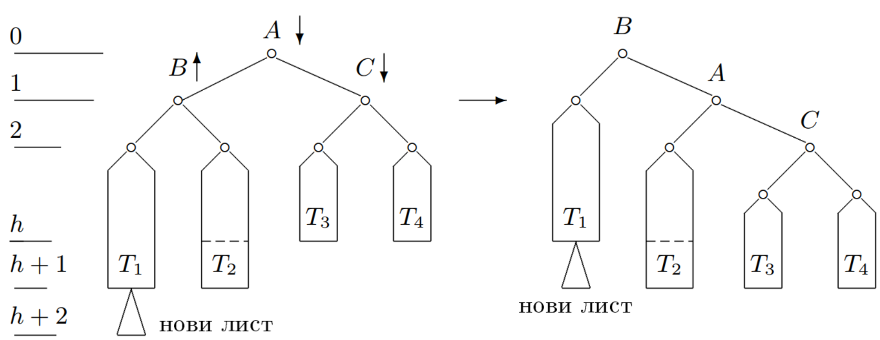
Primetimo da za svaki čvor AVL drveta važi da se visine levog i desnog poddrveta pre umetanja mogu razlikovati najviše za jedan. Nakon umetanja elementa u jedno od poddrveta, visine se mogu razlikovati najviše za dva.




Slika 5: Umetanja koja remete AVL svojstvo drveta.

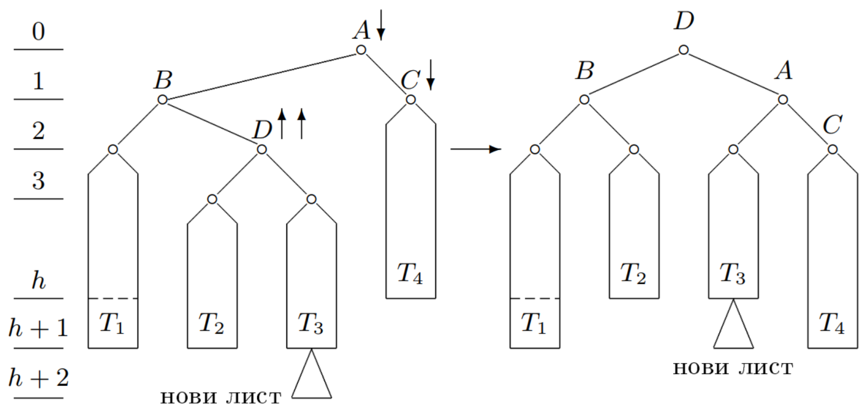
Pretpostavimo da je faktor ravnoteže u kritičnom čvoru $A$ jednak 1 (slika 5), odnosno da levo poddrvo drveta sa korenom u kritičnom čvoru ima za jedan veću visinu od desnog. Novi čvor $N$ može da završi u:

- desnom poddrvetu drveta sa korenom u kritičnom čvoru – u tom slučaju poddrvo kome je koren kritični čvor ostaje AVL i nije potrebno sprovoditi nikakve dodatne operacije.
- levom poddrvetu drveta sa korenom u kritičnom čvoru – u tom slučaju drvo prestaje da bude AVL jer visina levog drveta postaje za 2 veća od visine desnog i potrebno je intervenisati. Prema tome da li je novi čvor dodat levom ili desnom poddrvetu levog poddrveta, razlikujemo dva slučaja.
  - u slučaju kada je novi čvor dodat levom poddrvetu levog poddrveta na drvo se primenjuje tzv. *LL rotacija* (ilustrovana na slici 6): koren levog poddrveta $B$ drveta sa korenom u kritičnom čvoru $A$ „podiže" se za jedan nivo i postaje koren poddrveta kome je koren bio kritičan čvor, a ostatak drveta preuređuje se tako da drvo i dalje ostane uređeno binarno drvo. Drvo $T_1$ se podiže za jedan nivo ostajući i dalje levo poddrvo čvora $B$; drvo $T_2$ ostaje na istom nivou, ali umesto desnog poddrveta čvora $B$ postaje levo poddrvo čvora $A$ (primetimo da se na ovaj način uslov uređenosti drveta ne remeti); desno poddrvo čvora $A$ spušta se za jedan nivo. Pošto je $A$ kritičan čvor, faktor ravnoteže čvora $B$ je nužno 0, pa drveta $T_1$ i $T_2$ imaju istu visinu.[2] Novi koren poddrveta postaje čvor $B$, sa faktorom ravnoteže 0.
  - slučaj kada je novi čvor dodat desnom poddrvetu levog poddrveta je komplikovaniji; tada se drvo može uravnotežiti *LR rotacijom*, odnosno *dvostrukom rotacijom* (ilustrovanoj na slici 7). Novi čvor poddrveta umesto kritičnog čvora postaje desno dete levog deteta kritičnog čvora. Primetimo opet da pošto je $A$ kritični čvor, drveta $T_1$, $T_2$ i $T_3$ moraju pre umetanja imati istu visinu. Nakon LR rotacije, faktori ravnoteže čvorova $B$, $A$ i $D$ biće redom jednaki 1, 0 i 0, odnosno drvo postaje AVL.




Slika 6: LL rotacija.

[2] Drveta $T_3$ i $T_4$ ne moraju imati istu visinu, jer čvor $C$ nije na putu od kritičnog čvora do mesta umetanja.



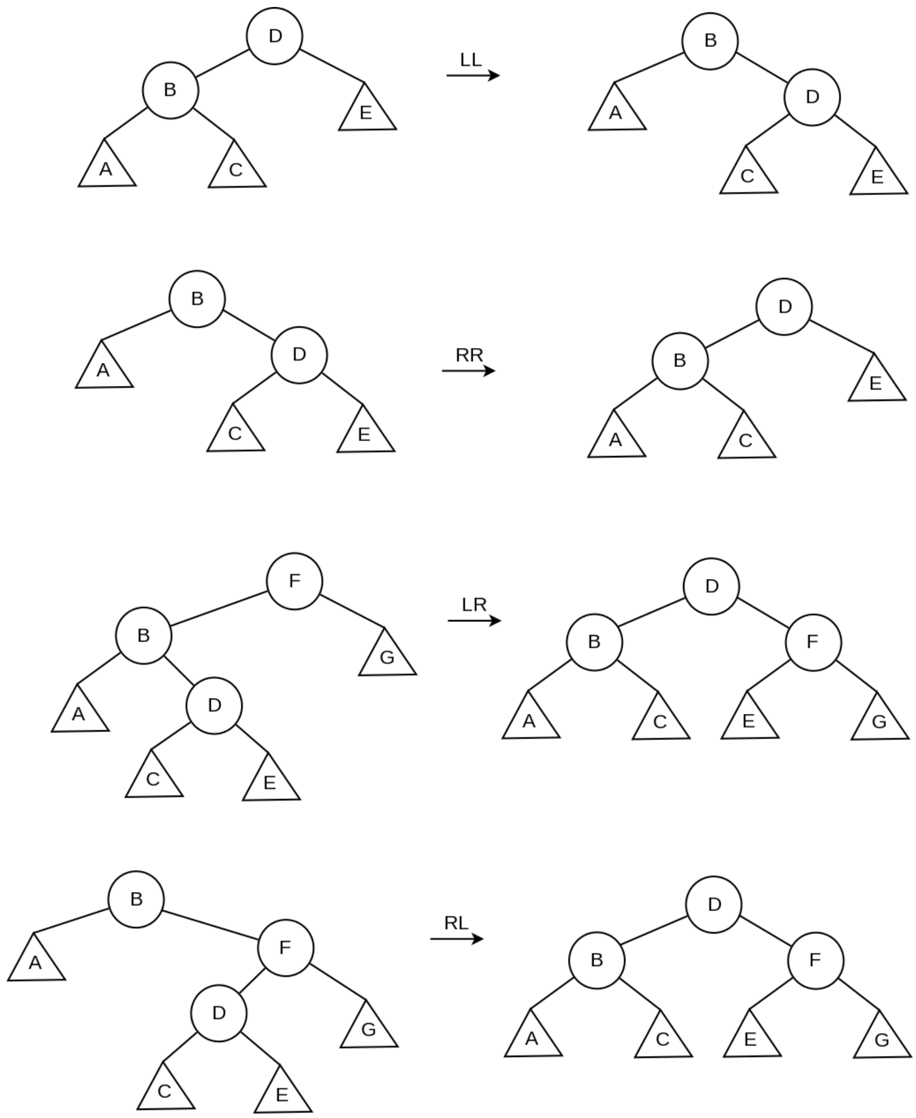

Slika 7: LR rotacija, odnosno dvostruka rotacija.

Simetrično, u slučaju kada je faktor ravnoteže u kritičnom čvoru jednak $-1$ (odnosno, kada je visina desnog poddrveta za jedan veća od visine levog), i kada se vrši umetanje čvora u desno poddrvo desnog poddrveta drveta sa korenom u kritičnom čvoru, potrebno je primeniti *RR rotaciju*, a ako se vrši umetanje u levo poddrvo desnog poddrveta uravnotežavanje se vrši *RL rotacijom*, odnosno dvostrukom rotacijom. Čitaocu ostavljamo da se uveri da se svim vrstama rotacija održavaju svojstva uređenosti drveta.

Sve pomenute vrste rotacija ilustrovane su na slici 8. Primetimo da kod LL i RR rotacije pre vršenja rotacije, a i nakon vršenja rotacija važi naredni odnos vrednosti čvorova $A < B < C < D < E$, dok se kod LR i RL rotacija takođe nakon rotacije čuva odnos vrednosti čvorova i on iznosi: $A < B < C < D < E < F < G$.

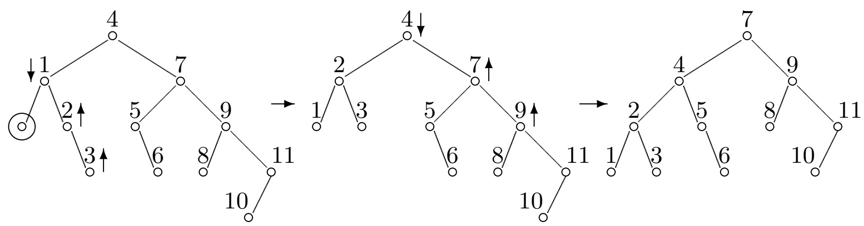
U implementaciji AVL drveta uz svaki čvor drveta čuva se njegova visina i njegov faktor ravnoteže, jednak razlici visina njegovog levog i desnog poddrveta; za AVL drvo su te razlike elementi skupa $\{-1, 0, 1\}$.

Zapazimo da u svakom od razmatranih slučajeva umetanja važi da *visina poddrveta kome je koren kritični čvor posle balansiranja ostaje ista kao što je bila pre umetanja novog elementa*. Odavde sledi da nakon balansiranja poddrveta kome je koren kritičan čvor nije potrebno sprovoditi nikakve dalje korekcije u drvetu (što, kao što ćemo videti, ne mora biti slučaj kada razmatramo operaciju brisanja iz AVL drveta). Ovo nam omogućava da se pri inicijalnom prolasku kroz drvo prilikom traženja mesta za umetanje čvora zapamti kritični čvor (poslednji čvor čiji je faktor ravnoteže različit od 0), i zatim se počev od kritičnog čvora ponovo prolazi put do umetnutog čvora i ažuriraju se vrednosti visina i faktora ravnoteže.




Slika 8: Ilustracija četiri vrste rotacija: LL rotacije, RR rotacije, LR rotacije i RL rotacije. $A$, $C$, $E$ i $G$ su balansirana uređena binarna drveta.

**Brisanje elementa iz AVL drveta** Brisanje iz AVL drveta je komplikovanije, kao i kod običnog uređenog binarnog drveta. U opštem slučaju se balansiranje drveta ne može izvesti pomoću samo jedne ili dve rotacije. Na primer, da bi se Fibonačijevo drvo $F_h$ sa $n$ čvorova uravnotežilo posle brisanja „loše“ izabranog čvora (npr. najvišeg lista), potrebno je izvršiti $h-2$, odnosno $O(\log n)$ rotacija (slika 9). U opštem slučaju je maksimalan potreban broj rotacija $O(\log n)$. Na sreću, svaka rotacija zahteva konstantni broj koraka, pa je vreme izvršavanja brisanja elementa iz AVL drveta, zajedno sa njegovim balansiranjem nakon umetanja, takođe ograničeno odozgo sa $O(\log n)$.




Slika 9: Uravnotežavanje Fibonačijevog drveta $T_4$ posle brisanja čvora, pomoću dve rotacije.

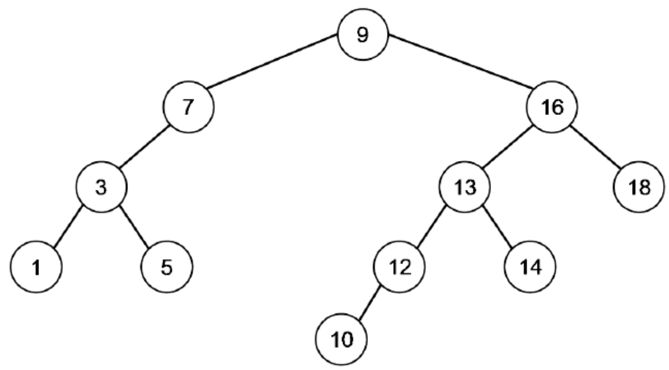
**Primene AVL drveta** AVL drveta predstavljaju vrstu samobalansirajućih uređenih binarnih drveta, koja obezbeđuju da im je visina u najgorem slučaju logaritamska u funkciji broja elemenata u drvetu. Odavde sledi da je složenost svake od operacija pretrage, umetanja i brisanja u najgorem slučaju složenosti $O(\log n)$.

AVL drveta su najstarija samobalansirajuća uređena binarna drveta. Nastala su iz potrebe efikasnije pretrage i dohvatanja podataka u bazama podataka, indeksiranjem podataka korišćenjem balansiranih binarnih drveta. AVL drveta su prevashodno korisna za organizaciju i čuvanje velike količnine podataka kada se operacije pretrage podataka sprovode često, dok su umetanja i brisanja ređa. Na primer, baza podataka koja bi čuvala informacije o postojećim linijama metroa u nekom gradu bi mogla da koristi AVL drvo za indeksiranje podataka jer se nove linije metroa dodaju izuzetno retko, dok sa druge strane veliki broj ljudi svakodnevno vrši pretragu baze u cilju rezervisanja karata.

AVL drveta se koriste i u svim drugim primenama koje zahtevaju efikasnu pretragu: mogu se koristiti za implementaciju raznih tipova internih kolekcija, kao što su uređeni skupovi i uređene mape. AVL drveta omogućavaju efikasno indeksiranje podataka u pretraživačima. Koriste se u fajl sistemima za čuvanje strukture direktorijuma i za implementaciju tabela simbola u kompajlerima i interpreterima (gde se uz identifikator čuvaju i pridružene informacije: tip podataka, doseg i sl). Takođe, mogu imati primenu u sistemima koji rade u realnom vremenu, kao što su sistemi za kontrolu letova ili u video igrama, kada je potrebno brzo sprovesti operacije nad podacima.

**Zadaci za vežbu**

1. Za koje čvorove datog drveta ne važi uslov balansiranosti?




2. Ilustrovati sve rotacije koje je potrebno izvršiti prilikom umetanja elemenata $1, 10, 5, 7, 3, 13, 6, 4, 8, 9$ u datom redosledu u prazno AVL drvo kako bi ono ostalo AVL.

3. Ilustrovati proces umetanja elemenata $1, 2, 3, 4, 5, 6, 7, 8$ u datom redosledu u prazno AVL drvo. Prikazati sve rotacije koje je potrebno izvršiti tokom umetanja elemenata kako bi drvo ostalo AVL.

4. Rotacije koje se vrše kod AVL drveta sa $n$ elemenata su složenosti:

   a. $\Theta(1)$

   b. $\Theta(\log n)$

   c. $\Theta(n)$

5. Pokazati da je dvostruka rotacija (LR i RL) kompozicija dve rotacije tipa LL ili RR.

6. U prazno AVL drvo se unose ključevi $6, 20, 45, 80$ i $96$. Utvrditi redosled umetanja koji će prouzrokovati dvostruku rotaciju (LR ili RL).

7. Neka su $a$, $c$ i $e$ proizvoljni čvorovi u poddrvetima $A$, $C$ i $E$ sa slike 8. Kako se menjaju visine ovih čvorova nakon svake od rotacija?

8. Koji je najmanji, a koji najveći broj čvorova koje može imati AVL drvo visine $h$?

9. Za koliko se mogu razlikovati visine dva lista u AVL drvetu od $n$ elemenata?

   a. najviše 1

   b. $O(\log n)$

   c. $O(n)$

10. Koja je maksimalna visina AVL drveta od 7 čvorova?