Teorija izračunljivosti

Ivan Prokić
kabinet 117, F blok
prokic@uns.ac.rs
http://imft.ftn.uns.ac.rs/∼iprokic/

Novi Sad

Tema 1

Sistemi za obračunavanje složenosti

Euklidov algoritam

Problem: Naći $NZD(a, b)$. Pretpostavimo da je $a \geq b > 0$.

Algorithm 1: Euklid (a, b)

$c = b$
while $c \neq 0$ **do**
  $c = a \, mod \, b$
  $a = b$
  $b = c$
return a

- Broj iteracija **while** petlje ograničen je sa $O(\log a)$ (primetimo: broj a ima približno $\log a$ cifara).

Pokazujemo da za ostatke koje dobijamo u iteracijama važi $r_{i+2} < \frac{r_i}{2}$. Imamo dva slučaja:
1. Ako je $r_{i+1} \leq \frac{r_i}{2}$, tvrđenje sledi iz $r_{i+2} < r_{i+1}$.
2. Ako je $r_{i+1} > \frac{r_i}{2}$, tada važi $1 < \frac{r_i}{r_{i+1}} < 2$, pa iz $r_i = q_{i+2} r_{i+1} + r_{i+2}$, tj. $\frac{r_i}{r_{i+1}} = q_{i+2} + \frac{r_{i+2}}{r_{i+1}}$, sledi $q_{i+2} = 1$ i $r_i = r_{i+1} + r_{i+2}$. Odavde je $r_{i+2} = r_i - r_{i+1} < \frac{r}{2}$.

Dakle, ako je $NZD(a, b) = r_{2n+1}$ imamo $a > b > r_1 > 2r_3 > \ldots > 2^n r_{2n+1} \geq 2^n$, odakle sledi $\lg a > n$. Zaključujemo da je broj iteracija ograničen sa $2 \lg a + 1 = O(\lg a) = O(\log a)$.

O veličini ulaza

Preostaje da procenimo vreme potrebno za jednu iteraciju **while** petlje. To zavisi od odabranog **sistema obračunavanja** vremenske složenosti osnovnih operacija:

- **univerzalni**: operacije se izvršavaju u konstantnom vremenu (što nije baš realno);
- **osnovni logaritamski**: vreme potrebno za operacije zavisi od veličine ulaza - ako je veći od argumenata n vreme potrebno za sve operacije je ograničeno sa $O(\log n)$;
- **strogi logaritamski**: kao osnovni logaritamski - ali za množenje i deljenje (realno) treba $O(\log^2 n)$.

Napomena: Sistem obračunavanja biramo prema posmatranom problemu - logaritmaski su precizniji ali često previše i nepotrebno komplikuju analizu.

Složenost Euklidovog algoritma:

- **univerzalni**: $T(n) = O(\log n)$;
- **osnovni logaritamski**: $T(n) = O(\log n) \cdot O(\log n) = O(\log^2 n)$;
- **strogi logaritamski**: $T(n) = O(\log n) \cdot O(\log^2 n) = O(\log^3 n)$.

Prost broj

Problem 1: Proveriti da li je $n > 1$ prost broj.

Algorithm 2: ProstBroj (n)

$bool = True$
$m = \lfloor \sqrt{n} \rfloor$
for $k = 2$ **to** m **do**
  if $n \ mod \ k = 0$ **then**
    $bool = False$
return $bool$

Složenost:

- **univerzalni**: $T(n) = O(\sqrt{n})$;
- **strogi logaritamski**: $T(n) = O(\sqrt{n} \log^2 n)$.

Kolika je složenost računanja $\lfloor \sqrt{n} \rfloor$?

Faktorijel

Problem 2: Za $n > 1$ izračunati faktorijel.

Algorithm 3: Faktorijel (n)

$f = 1$
for $i = 2$ **to** n **do**
  $f = f \cdot i$
return f

Složenost:

- **univerzalni**: $T(n) = O(n)$;
- **strogi logaritamski**: Kako je $(n-1)! < n^n$, broj cifara $(n-1)!$ možemo ograničiti sa $O(n \log n)$, pa je za množenje sa n, koji ima $O(\log)$ cifara, potrebno $O(n \log^2 n)$. S obzirom da petlja iterira n puta, imamo $T(n) = O(n^2 \log^2 n)$.

Ocena može da se poboljša ako se koristi neki efikasniji algoritam za množenje (npr. Strassen-ov).

Tema 2

Asimptotske granice za rekurentne relacije

Ocena složenosti rekurentne relacije

Za ocenu složenosti rekurentne relacije, tj. određivanje asimptotskog ograničenja možemo koristiti sledeće metode:

- **Metod substitucije**: pogađamo rešenje i koristimo matematičku indukciju da dokažemo da smo pogodili;
- **Rekurzivna stabla**: konvertujemo rekurentnu relaciju u stablo čiji svaki čvor predstavlja cenu za odgovarajući nivo rekurzivnog poziva;
- **Master metoda**: daje asimptotsko ograničenje za rekurentne relacije oblika
$$T(n) = aT(n/b) + f(n),$$
gde su $a \geq 1, b > 1$ i $f(n)$ je data funkcija. Ovakve rekurentne relacije se pojavljuju u divide and conquer algoritmima.
- **Linerane rekurentne relacije**: homogene i nehomogene.

Napomena: neke rekurentne relacije se mogu rešiti direktnom zamenom *teleskopiranjem*

Tema 3

Metod substitucije

Hanojska kula

Primer (Hanojska kula)

Na jednom štapu naslagano je n diskova različite veličine, poslaganih po veličini: najveći na dnu, najmanji na vrhu. Pored tog štapa postoje još dva. Cilj je premestiti sve diskove sa jednog štapa na drugi, tako da se u svakom koraku premešta samo jedan disk sa vrha jedne gomile i da nijedan disk nikada ne sme biti smešten na manji. Koliko koraka je potrebno?

Rešenje. Ako je $n = 1$ potreban je samo jedan korak, tj. $T(1) = 1$. Za $n > 1$, problem rešavamo rekurzivno, premestimo $n - 1$ diskova na drugi štap, zatim najveći disk na treći, i na kraju sa drugog premestiomo $n - 1$ diskova na treći štap. Dakle, za $n > 1$ vredi

$$T(n) = 2T(n - 1) + 1.$$

◇

Hanojska kula

Rekurentnu relaciju rešavamo teleskopiranjem:

$$
\begin{aligned}
T(n) &= 2T(n-1)+1 \\
&= 2(2T(n-2)+1)+1 \;=\; 2^2T(n-2)+2+1 \\
&= 2^2(2T(n-3)+1)+2+1 \;=\; 2^3T(n-3)+2^2+2+1 \\
&= \dots \\
&= 2^{n-1}T(1)+\sum_{i=0}^{n-2}2^i \\
&= 2^{n-1}+\frac{2^{n-1}-1}{2-1} \\
&= 2^n-1=\Theta(2^n).
\end{aligned}
$$

Merge sort

Za MergeSort smo imali $T(n) = 2T(n/2) + \Theta(n)$ i rekli smo da je $T(n) = \Theta(n \lg n)$. Ako uvedemo smenu $n = 2^m$ ovu ocenu složenosti možemo izvesti teleskopiranjem. Prvo ćemo $\Theta(n)$ zapisati kao cn.

$$\begin{aligned}
T(2^m) &= 2T(2^{m-1}) + c2^m \\
&= 2(2T(2^{m-2}) + c2^{m-1}) + c2^m \;=\; 2^2T(2^{m-2}) + c2^m \cdot 2 \\
&= 2^2(2T(2^{m-3}) + c2^{m-2}) + c2^m \cdot 2 \;=\; 2^3T(2^{m-3}) + c2^m \cdot 3 \\
&= \ldots \\
&= 2^mT(1) + c2^m m.
\end{aligned}$$

Vraćanjem smene $m = \lg n$ i pošto je $T(1)$ konstanta dobijamo

$$T(n) = nT(1) + cn \lg n = \Theta(n \lg n).$$

Videćemo kasnije da ovu ocenu možemo izvesti i pomoću rekurzivnog stabla i Master metode.

Merge sort

Napomena: Ne postoji generalni način kojim se može pogoditi ocena svake rekurentne relacije.

Pošto smo za Merge sort uspeli da damo ocenu $T(n) = \Theta(n \lg n)$, sada možemo matematičkom indukcijom da dokažemo da smo u pravu. Zapravo, treba da pokažemo:

- $T(n) = O(n \lg n)$, tj. $T(n) \leq dn \lg n$, za odgovarajuće $d > 0$;
- $T(n) = \Omega(n \lg n)$, tj. $T(n) \geq en \lg n$, za odgovarajuće $e > 0$;

Pokažimo da za $T(n) = T(\lfloor n/2 \rfloor) + T(\lceil n/2 \rceil) + cn$ važi $T(n) \leq cn \lg n$.

Indukcijska hipoteza: Pretpostavimo da nejednakost važi za sve $m < n$, pa i za $m = \lfloor n/2 \rfloor$ i $m = \lceil n/2 \rceil$, odnosno

$$T(\lfloor n/2 \rfloor) \leq d\lfloor n/2 \rfloor \lg \lfloor n/2 \rfloor \quad \text{i} \quad T(\lceil n/2 \rceil) \leq d\lceil n/2 \rceil \lg \lceil n/2 \rceil.$$

Merge sort

Induktivni korak:

$$
\begin{aligned}
T(n) &= T(\lfloor n/2 \rfloor) + T(\lceil n/2 \rceil) + cn \\
&\leq d\lfloor n/2 \rfloor \lg \lfloor n/2 \rfloor + d\lceil n/2 \rceil \lg \lceil n/2 \rceil + cn \\
&\leq d\frac{n}{2} \lg \frac{n}{2} + d\frac{n+1}{2} \lg \frac{n+1}{2} + cn \\
&\leq \frac{dn}{2}(\lg n - 1) + \frac{d(n+1)}{2} \lg n + cn \\
&\leq dn \lg n - \frac{dn}{2} + \frac{dn}{4} + cn \leq dn \lg n,
\end{aligned}
$$

za $d \geq 4c$, gde smo koristili $\lfloor n/2 \rfloor \leq \frac{n}{2}$, $\lceil n/2 \rceil \leq \frac{n+1}{2}$, $\frac{n+1}{2} \leq n$ i $\lg n \leq \frac{n}{2}$ (za $n \geq 1$).

Merge sort

Baza indukcije: Treba pokazati $T(1) \leq d \lg 1 = 0$?

Međutim, $T(n) \leq dn \lg n$, po definiciji O-notacije treba da važi za sve $n \geq n_0$, gde konstantu n_0 mi možemo da biramo. Ako u našem primeru uzmemo $n_0 = 2$, bazu indukcije možemo sprovesti za $T(2)$ i $T(3)$.

Baza indukcije: Ako je $T(1) = c$, imamo da bazni slučajevi

$$\begin{aligned} T(2) &= 2T(1) + 2c = 4c \leq 2d \lg 2 = 2d \\ T(3) &= T(1) + T(2) + 3c = 8c \leq 3d \lg 3, \end{aligned}$$

važe za $d \geq 4c$.

Tema 4

Rekurzivna stabla

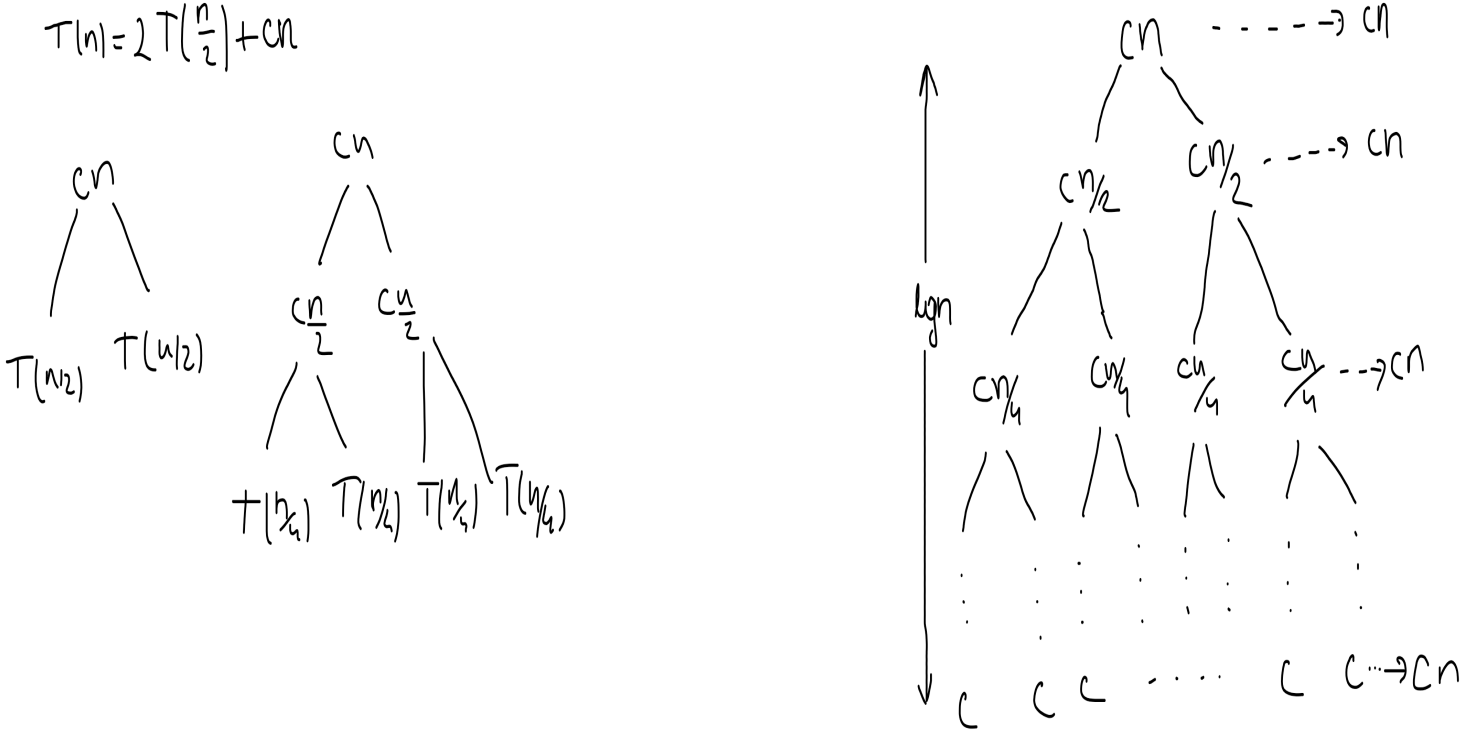
Rekurzivna stabla

Rekurzivna stabla možemo koristiti na dva načina:

- za dobijanje ocene, koju posle verifikujemo metodom supstitucije;
- za direktno rešenje rekurentne relacije (ako smo dovoljno precizni).

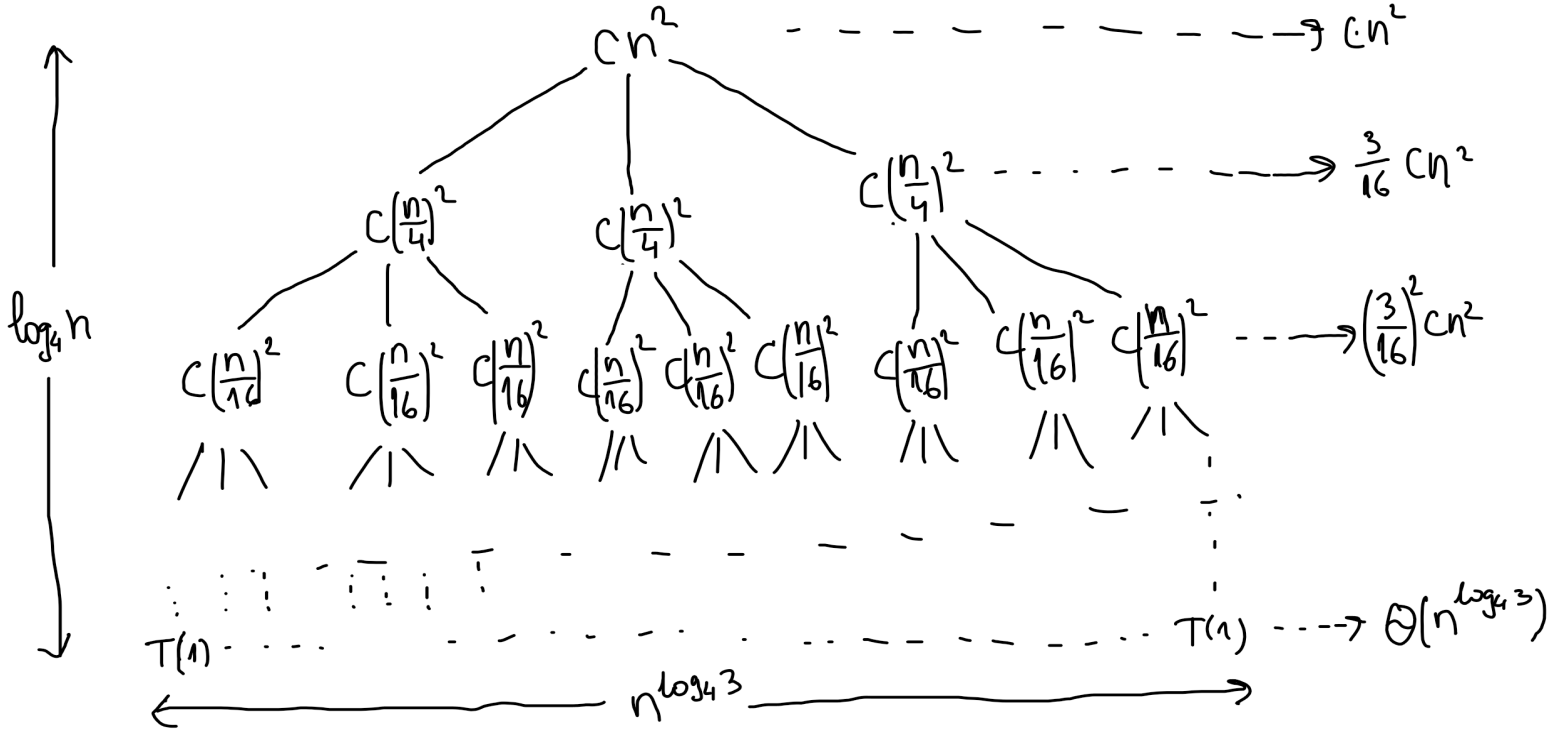
Merge sort

Neka je ponovo $n = 2^m$. Iz rekurzivnog stabla za MergeSort rekurentnu relaciju vidimo: cena za svaki nivo (ukupno) je cn (piše desno); visina stabla je $\lg n$ (piše levo), pa je broj nivoa $\lg n + 1$. Dakle, $T(n) = cn \lg n + cn = \Theta(n \lg n)$, a dokaz da je ovo tačno smo već izveli koristeći metod supstitucije.



Primer: $T(n) = 3T(\lfloor n/4 \rfloor) + \Theta(n^2)$

Dati ocenu za rekurentnu relaciju $T(n) = 3T(\lfloor n/4 \rfloor) + \Theta(n^2)$. Prvo ćemo uzeti da je $n = 4^m$ i rekurziju ćemo zapisati u obliku $T(n) = 3T(n/4) + cn^2$, gde je $c > 0$ (jer izvodimo samo ocenu koju ćemo kasnije verifikovati metodom supstitucije).



Primer: $T(n) = 3T(\lfloor n/4 \rfloor) + \Theta(n^2)$

Na i-tom nivou:

- veličina argumenta u čvoru je $\frac{n}{4^i}$ - pa je na posledenjem nivou $\frac{n}{4^i} = 1$, odakle je $i = \log_4 n$, a broj nivoa je $\log_4 n + 1$;
- broj čvorova je 3^i i svaki čvor vredi $c\left(\frac{n}{4^i}\right)^2$- pa je cena za čitav nivo $cn^2\left(\frac{3}{16}\right)^i$. Na poslednjem nivou ima $3^{\log_4 n} = n^{\log_4 3}$ čvorova. Kako je $T(1)$ konstana cena poslednjeg nivoa je $\Theta(n^{\log_4 3})$.

$$
\begin{aligned}
T(n) &= \sum_{i=0}^{\log_4 n - 1} cn^2 \left(\frac{3}{16}\right)^i + \Theta(n^{\log_4 3}) \\
&< cn^2 \sum_{i=0}^{\infty} \left(\frac{3}{16}\right)^i + \Theta(n^{\log_4 3}) \\
&= cn^2 \frac{1}{1 - \frac{3}{16}} + \Theta(n^{\log_4 3}) = O(n^2),
\end{aligned}
$$

jer je $\log_4 3 < 2$.

Primer: $T(n) = 3T(\lfloor n/4 \rfloor) + \Theta(n^2)$

Dakle, $T(n) = O(n^2)$ (gornja granica). Ako pogledamo koren stabla vidimo da takođe mora važiti $T(n) = \Omega(n^2)$ (donja granica), odakle zaključujemo da je $T(n) = \Theta(n^2)$ (tesna granica).

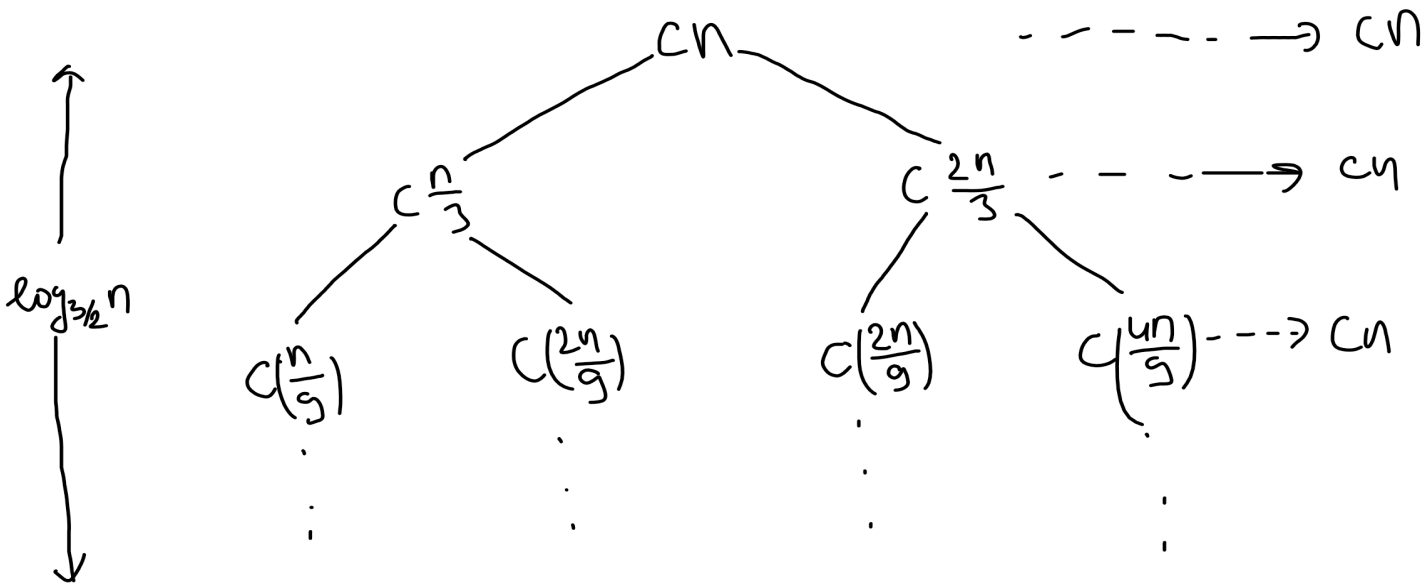
Koristeći metod substitucije dokazaćemo $T(n) = O(n^2)$, tj. $T(n) \leq dn^2$, za neku konstantu $d > 0$.

$$\begin{aligned} T(n) &= 3T(\lfloor n/4 \rfloor) + cn^2 \\ &\leq 3d\lfloor n/4 \rfloor^2 + cn^2 \\ &\leq 3d(n/4)^2 + cn^2 = \frac{3}{16}dn^2 + cn^2 \leq dn^2, \end{aligned}$$

za $d \geq \frac{16c}{13}$.

Primer: $T(n) = T(n/3) + T(2n/3) + O(n)$

Dati ocenu za rekurentnu relaciju $T(n) = T(n/3) + T(2n/3) + O(n)$.



Napomena: Od korena do listova nisu svi putevi iste dužine: put $n \to (1/3)n \to (1/9)n \to \ldots \to (1/3)^i n$ je najkraći, a $n \to (2/3)n \to (4/9)n \to \ldots \to (2/3)^i n$ najduži.

Primer: $T(n) = T(n/3) + T(2n/3) + O(n)$

(Pošto ćemo dati gornje ograničenje posmatramo najduži put.)
Na i-tom nivou:

- veličina argumenta u čvoru je (najviše) $n\left(\frac{2}{3}\right)^i$ - pa je na posledenjem nivou $n\left(\frac{2}{3}\right)^i = 1$, odakle je $i = \log_{3/2} n$, a broj nivoa je $\log_{3/2} n + 1$;
- broj čvorova je (najviše) 2^i;
- cena na svakom nivou je najviše cn: svaki čvor na $i-1$-om nivou (sa cenom x) je povezan sa 2 na i-tom (sa cenama $x/3$ i $2x/3$), pa je cena na svim nivoima manja ili jednaka ceni prvog nivoa.

Umesto da se upuštamo u tačno računanje složenosti, sada možemo pretpostaviti da je $T(n) = O(n \log_{3/2} n) = O(n \lg n)$ i, koristeći metod supstitucije, dokazati da smo u pravu.

Primer: $T(n) = T(n/3) + T(2n/3) + O(n)$

Dokažimo $T(n) = O(n \log_{3/2} n) = O(n \lg n)$, tj. $T(n) \leq dn \lg n$, za neku konstantu $d > 0$.

$$
\begin{aligned}
T(n) &= T(n/3) + T(2n/3) + cn \\
&\leq d\frac{n}{3} \lg \frac{n}{3} + d\frac{2n}{3} \lg \frac{2n}{3} + cn \\
&= d\frac{n}{3} \lg n + d\frac{2n}{3} \lg 2n - dn \lg 3 + cn \\
&= d\frac{n}{3} \lg n + d\frac{2n}{3} \lg n + \frac{2dn}{3} - dn \lg 3 + cn \\
&= dn \lg n + n\left(\frac{2d}{3} - d \lg 3 + c\right) \\
&\leq dn \lg n,
\end{aligned}
$$

za $d \geq \frac{c}{\lg 3 - 2/3}$.

Zadaci

Zadatak

Za sledeće rekurentne relacije koristeći rekurzivna stabla odrediti što bolju gornju granicu. Koristeći metod substitucije dokazati da su granice odgovarajuće.

1. $T(n) = T(n/3) + n^2$;
2. $T(n) = 2T(\lfloor n/3 \rfloor) + n$;
3. $T(n) = 3T(\lfloor n/3 \rfloor) + n$;
4. $T(n) = 4T(\lfloor n/3 \rfloor) + n$;
5. $T(n) = 4T(n/3 + 2) + n$;
6. $T(n) = 2T(n-1) + 1$;
7. $T(n) = T(n-1) + T(n/2) + n$;
8. $T(n) = 4T(\lfloor n/2 \rfloor) + 4n$.