14. Összehasonlításos rendezések alaptételei

Alsó korlát becslés az összehasonlító rendezésekre.
Vegyünk egy F feladatot, valamint A és b algoritmusokat.

Az A algoritmus jobb (nem rosszabb), mint a B algoritmus
$\Leftrightarrow MT_A(n) \leq MT_B(n)\ (\forall n \in természetes)$.

Az A algoritmus átlagosan jobb (nem rosszabb), mint a B algoritmus
$\Leftrightarrow AT_A(n) \leq AT_B(n)\ (\forall n \in természetes)$.

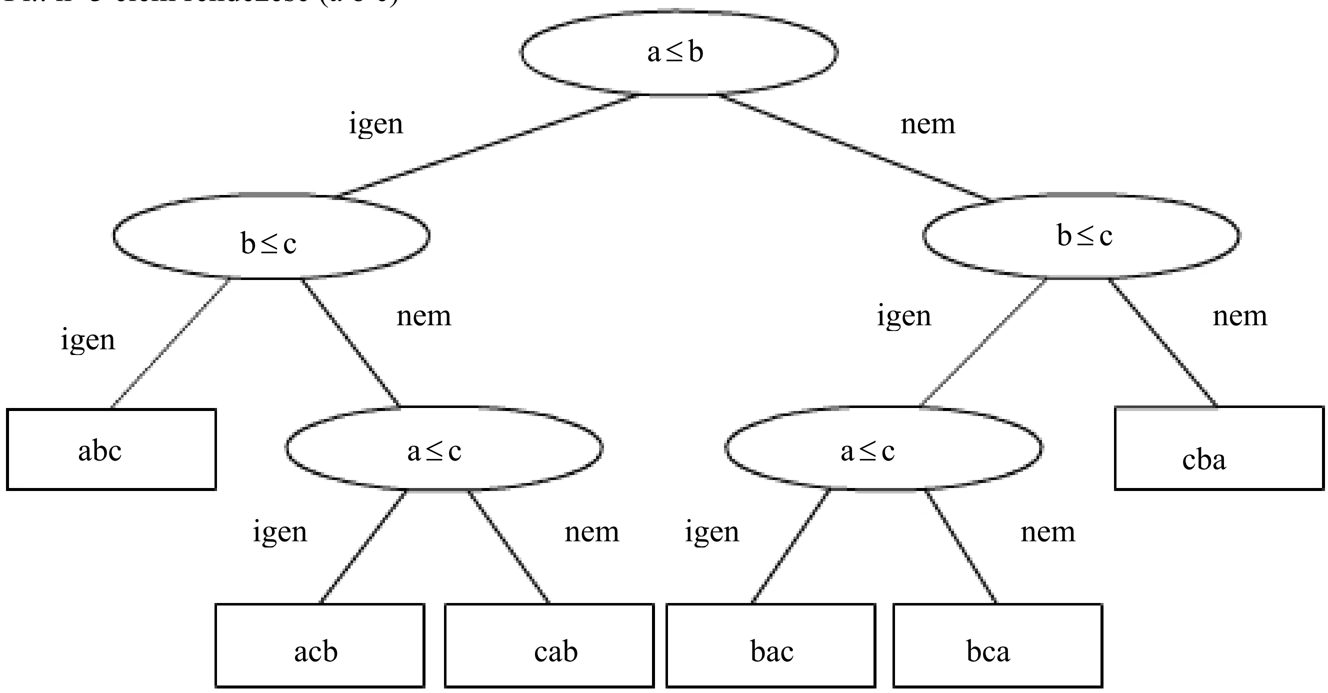
Az S algoritmus optimális $\Leftrightarrow MT_S(n) \leq MT_R(n)$ ($\forall$ elképzelhető algoritmusra).

Az S algoritmus átlagosan optimális $\Leftrightarrow AT_S(n) \leq AT_R(n)$ ($\forall$ elképzelhető algoritmusra).

$$MT_R(n) = \Omega(MT_S(n)) \qquad \textit{és} \qquad AT_R(n) = \Omega(AT_S(n))$$

Belátjuk, hogy
1) Nincs olyan összehasonlító rendezés, amely n hosszú sorozatra a legrosszabb esetben $n \log n$-nél nagyságrendben kevesebb összehasonlítást végezne.
   $MÖ_R(n) = \Omega(n \log n)$ ($\forall$ R összehasonlító-rendező algoritmusra).
2) Átlagosan is hasonló mondható
   $AÖ_R(n) = \Omega(n \log n)$ ($\forall$ R összehasonlító-rendező algoritmusra).

Döntési fa
Pl.: n=3 elem rendezése (a b c)



A döntési fa belső pontjaiban kérdések vannak, mindig két kimenettel (tökéletes fa).

Lemma: Bármely r összehasonlító-rendező algoritmus esetén $MÖ_R(n) \geq \log_2(n!)$.
Bizonyítás:

nézzük az R algoritmus döntési fáját n-re, $MÖ_R(n) = h(t)$

$h(t)$-nek elég nagynak kell lennie, hogy legyen a fában $n!$ levél, hiszen mind az $n!$ inputra működnie kell R-nek

pl.: n=3-ra $h(t)$=2 levél lenne, legalább $h(t)$=3 kell legyen!

Szükséges:

$$2^{h(t)} \geq (n!) \qquad (pl.: 2^3 \geq 3! = 6)$$

$$h(t) \geq \log_2(n!)$$

$MÖ_R(n) \geq \log_2(n!)$.

Megjegyzés: $n = 3$ $\log_2 6 = 2{,}6 \to 3\,összehasonlítás$

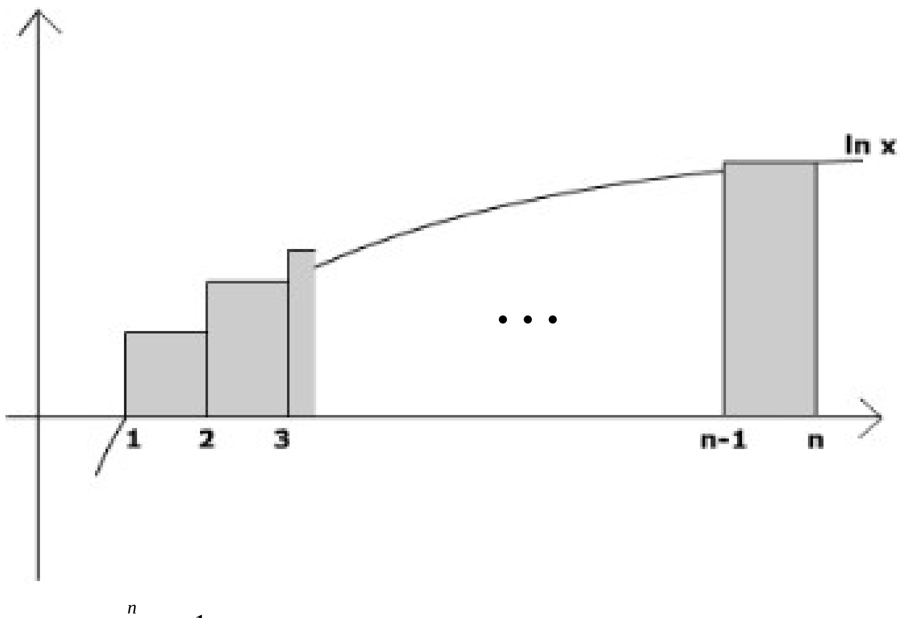
$n = 4$ $\log_2 24 = 4{,}6 \to 5\,összehasonlítás$

$n = 5$ $\log_2 120 = 6{,}9 \to 7\,összehasonlítás$

1. Tétel: $MÖ_R(n) = \Omega(n \log n)$ ($\forall n, \forall$ R összehasonlító rendezésre).

Bizonyítás:

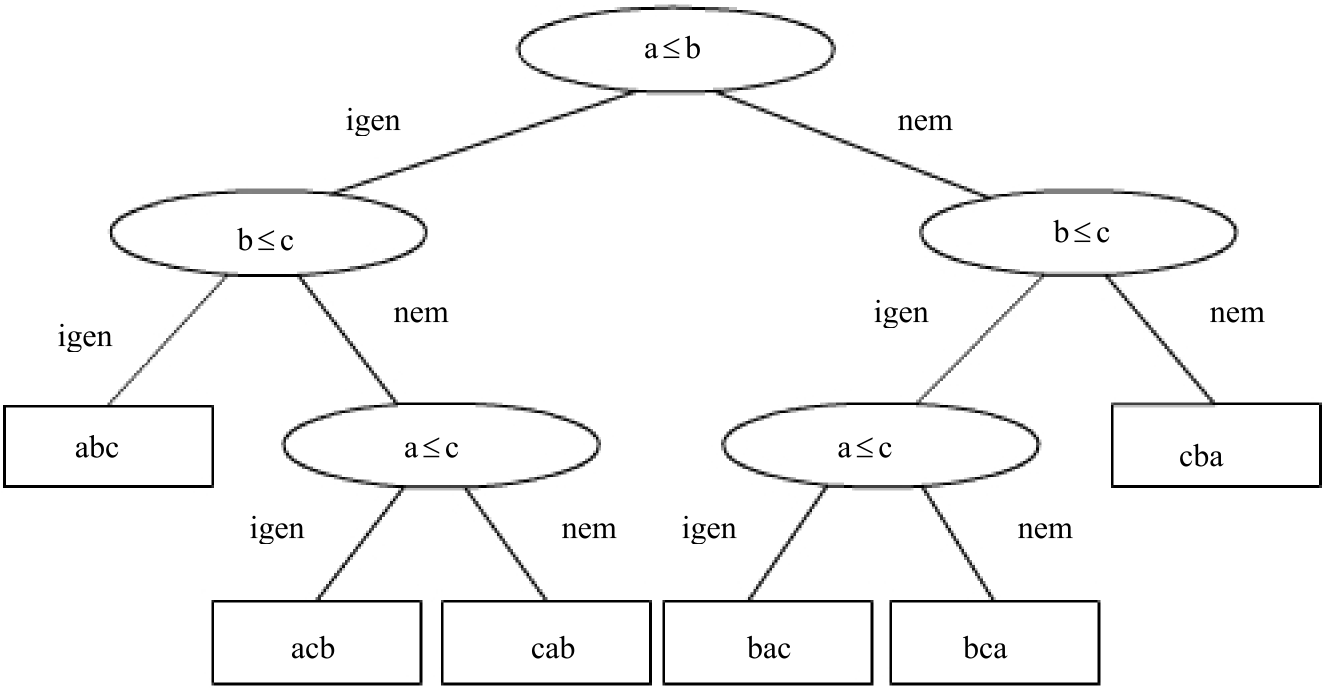
$$MÖ_R(n) \geq \log_2(n!) = \frac{1}{\ln 2}\ln(n(n-1)\cdot\ldots\cdot 2\cdot 1) = \frac{1}{\ln 2}(\ln 1 + \ldots + \ln n) >$$



$$> \frac{1}{\ln 2}\int_1^n \ln x dxx = \frac{1}{\ln 2}\left[x \ln x - x\right]_1^n = \frac{1}{\ln 2}(n \ln n - n + 1) \approx n \log_2 n - 1{,}44n + 1{,}44 = \Omega(n \log n)$$

$$\left(\begin{array}{l} \dfrac{1}{\ln 2} = 1{,}44; \\ n \log_2 n - 1{,}44n + 1{,}44 \approx \Theta(n \log n) \end{array}\right)$$

t=



$M\ddot{O}(n) \geq \log_2(n!) \qquad ,ahol\ M\ddot{O}(n) = h(t)$

$M\ddot{O}(n) = \Omega(n \log n)$

$h(t) \geq n \log n - 1{,}44n + 1{,}44 \approx \Theta(n \log n)$

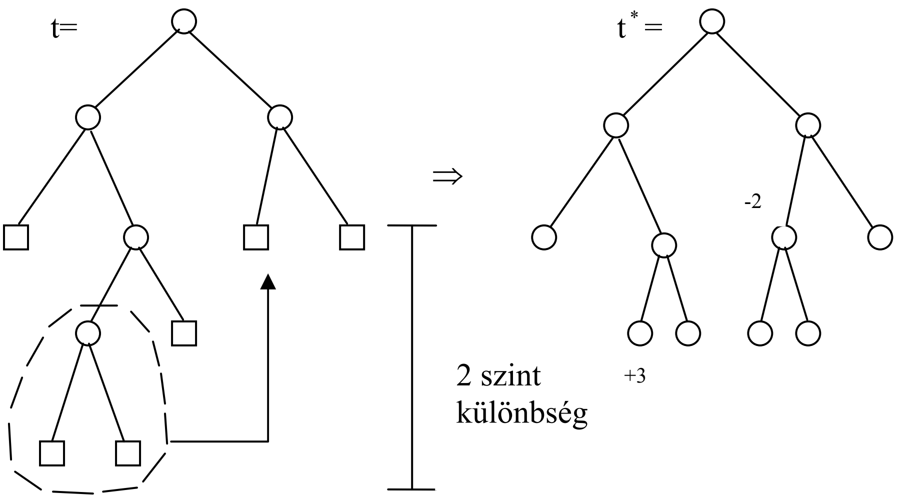
Átlagos összehasonlítás szám

Mi is ez pontosan? pl.: itt most $A\ddot{O}_t(3) = \dfrac{2+3+3+3+3+2}{6}$

Általában $A\ddot{O}_t(n) = \dfrac{tlevélmagasságainakazösszege}{n!} = \dfrac{lhsum(t)}{n!}$.

Az alsó korlát megadásaihoz a/egy „legjobb” döntési fát veszünk.

Állítás: A legjobb/optimális döntési fa majdnem teljes.
Bizonyítás:

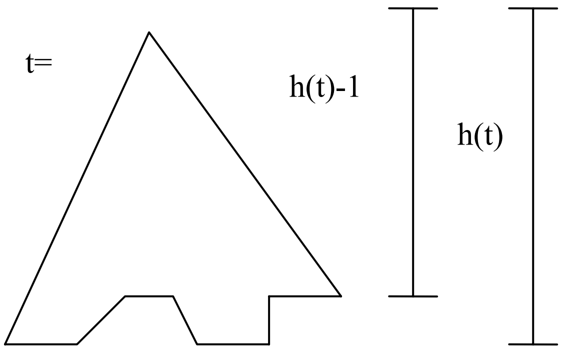


A számok a csúcstól való távolságot jelzik

lhsum(t)=2+4+4+3+2+2+=17
lhsum(t^*)=2+3+3+3+3+2=16

,azaz lhsum(t^*)=lhsum(t)+3-2-1-1=lhsum(t)-1

Optimális döntési fa



h(t) $\geq \log_2(n!)$

lhsum(t)>n!(h(t)-1)
$\geq n!\,(n\log_{2n} - 1{,}44n + 1{,}44 - 1)$

$$A\ddot{O}(n) = \frac{lhsum(t)}{n!} > \frac{1}{n!} n!(n\log_2 n - 1{,}44n) = \Omega(n\log n)$$

$$\left(n\log_2 n - 1{,}44n \approx \Theta(n\log n)\right)$$