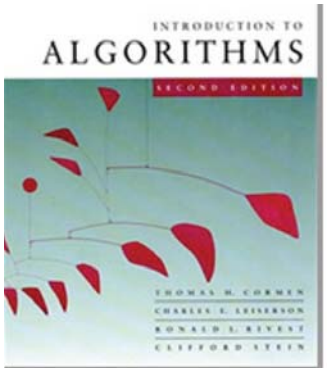
Algoritmalara Giriş

6.046J/18.401J

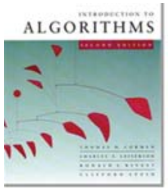


DERS 12

Atlama Listeleri

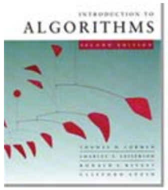
- Veri Yapısı
- Rastgele Araya Yerleştirme
- "Yüksek olasılıkla" sınırı
- Analiz (Çözümleme)
- Yazı Tura Atma

Prof. Erik D. Demaine



Atlama Listeleri

- Basit Rastgele Dinamik Arama Yapısı
  - William Pugh tarafından 1989'da geliştirildi.
  - Uygulaması kolay
- n elemanlı dinamik bir kümeyi, her bir işlem için $O(\lg n)$ zamanlı beklenen ve yüksek ihtimalle kurar.
  - $T(n)$'nin kuyruk dağılımında sıkı garanti.
  - "Hemen her zaman" $O(\lg n)$

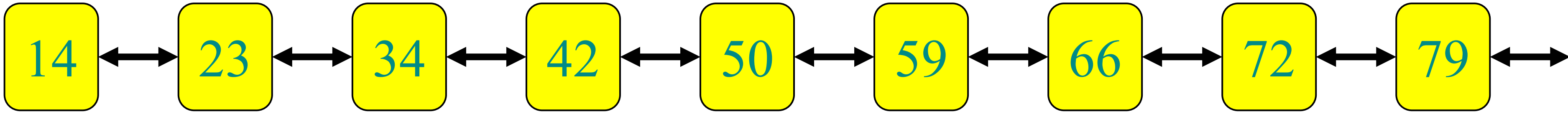


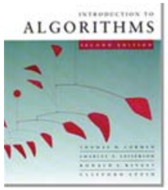
Bir Bağlantılı Liste

En basit veri yapısından başlangıç:

(sıralı) bağlantılı liste

- Aramalar en kötü durumda $\Theta(n)$ kadar zaman alır.
- Aramaları nasıl hızlandırabiliriz?

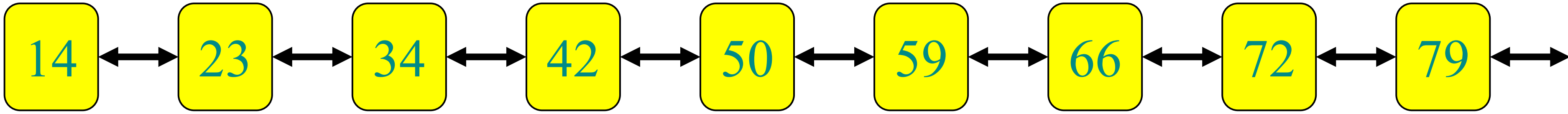


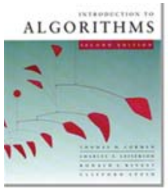


İki Bağlantılı Liste

İki sıralı bağlantılı listemiz olduğunu düşünün.
(Elemanların alt kümeleriyle)

- Her eleman bir veya iki listede yer alabilir.
- Aramaları nasıl hızlandırabiliriz?

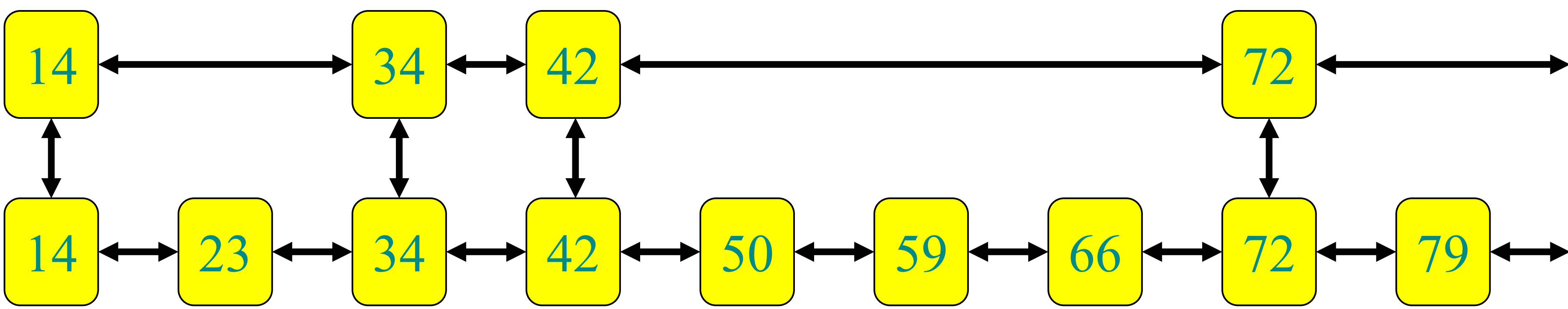


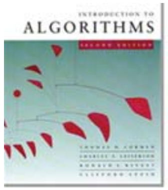


Metro Gibi İki Bağlantılı Liste

Fikir: Ekspres ve yerel Metro Hatları
(New York 7. cadde Hattı)

- Ekspres hat birkaç durağı birbirine bağlar.
- Yerel hat bütün durakları birbirine bağlar.
- Ana istasyonlar arasını bağlantılar. (linkler)

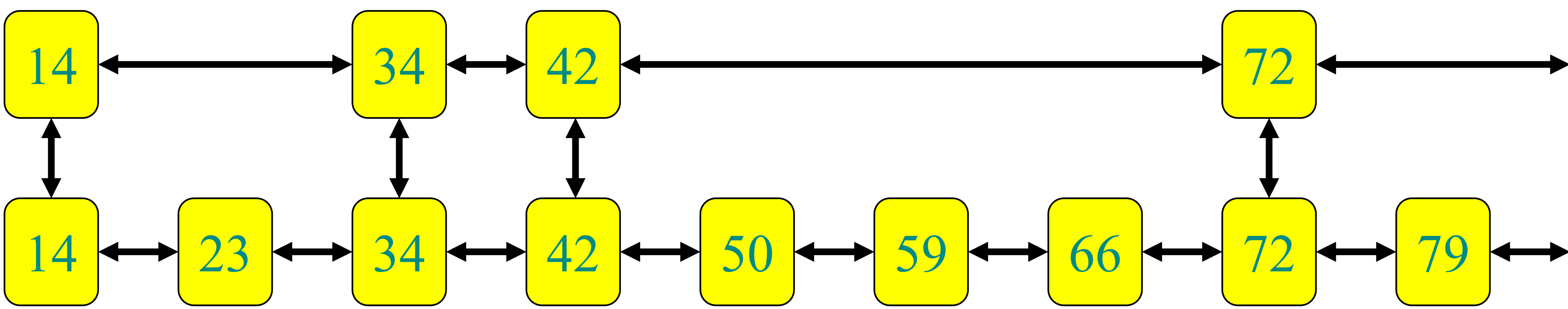


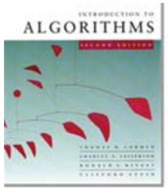


İki Bağlantılı Listede Arama

SEARCH / ARA (x) :

- Üst bağlantılı listede (L_1) sağa doğru çok uzak olana kadar yürüyün.
- Alt bağlantılı listeye (L_2) yürüyün.
- L_2'de eleman bulunana veya bulunamayana kadar yürüyün.



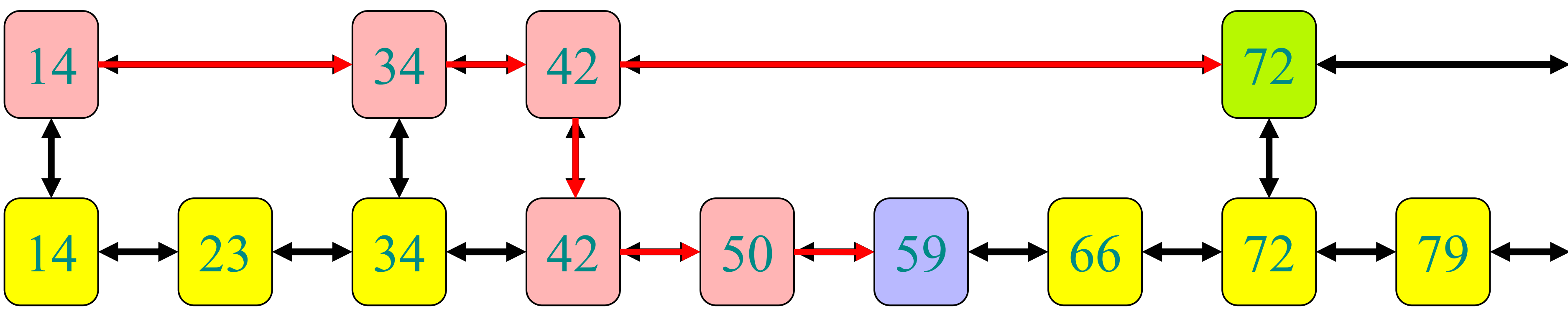


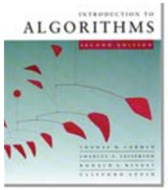
İki Bağlantılı Listede Arama

Örnek: SEARCH(59)

Çok uzak:

$59 < 72$

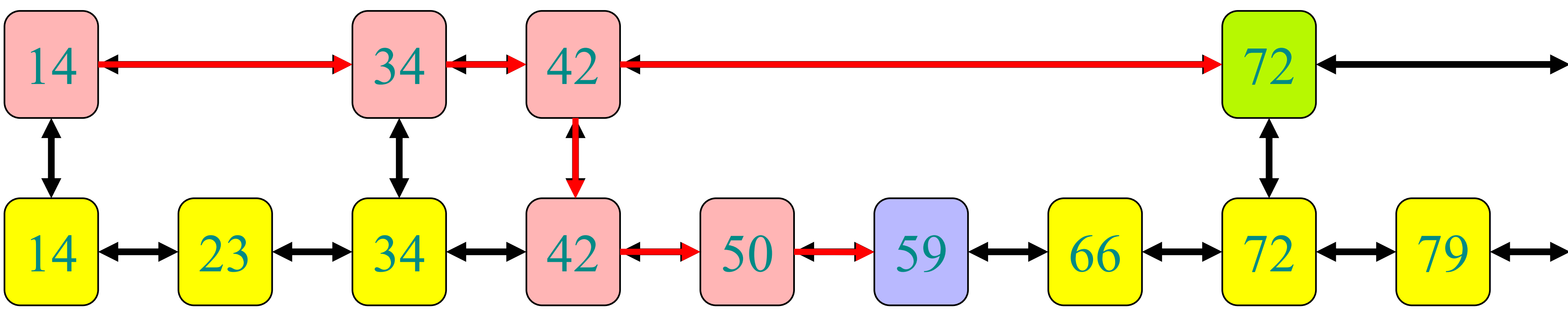


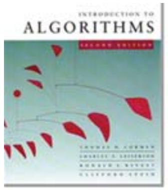


İki Bağlantılı Listenin Tasarımı

Soru : Hangi düğümler L_1'de olmalı?

- Bir metroda, "popüler istasyonlar",
- Bizi burada en kötü durum performansı ilgilendiriyor.
- **En iyi yaklaşım :** L_1'deki düğümleri eşit mesafede yerleştirin.
- Peki, L_1'de kaç tane düğüm olmalı?





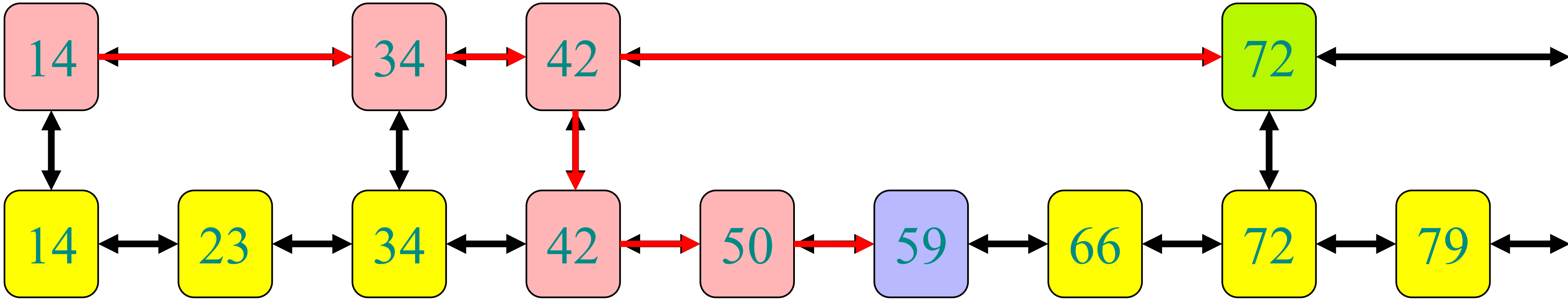
İki Bağlantılı Listenin Analizi

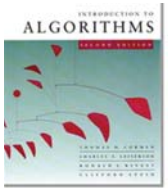
Çözümleme :

- Arama maliyeti kabaca $|L_1| + \dfrac{|L_2|}{|L_1|}$

- Terimler eşit iken (sabit katlara kadar) azalır.

- $|L_1|^2 = |L_2| = n \Rightarrow |L_1| = \sqrt{n}$



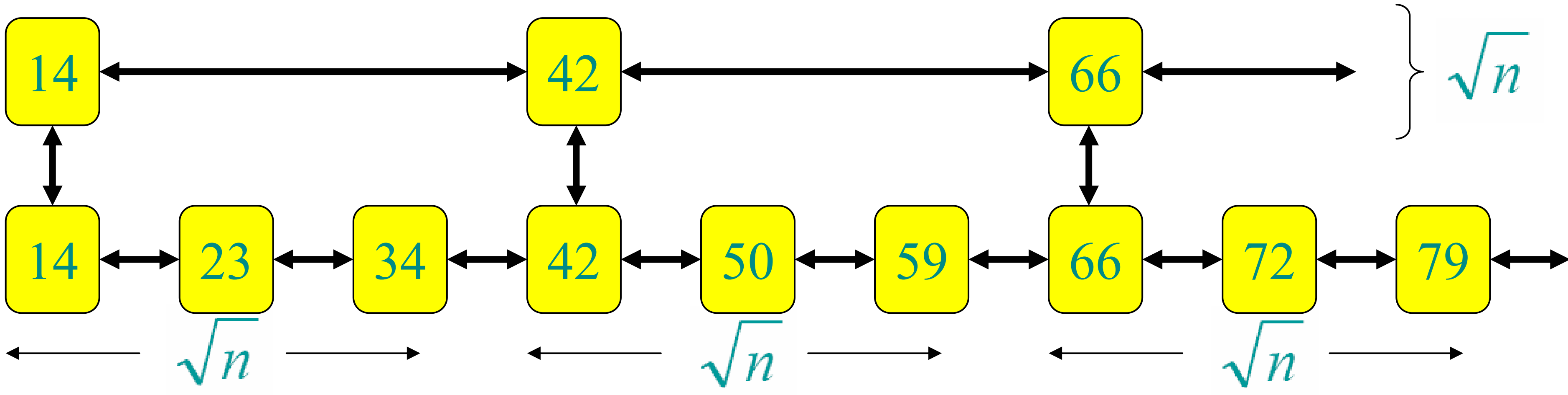


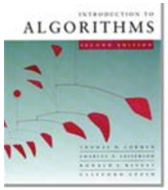
İki Bağlantılı Listenin Analizi

¢ º z¿mleme:

- $|L_1| = \sqrt{n}$, $|L_2| = n$
- Arama maliyeti kabaca

$$|L_1| + \frac{|L_2|}{|L_1|} = \sqrt{n} + \frac{n}{\sqrt{n}} = 2\sqrt{n}$$



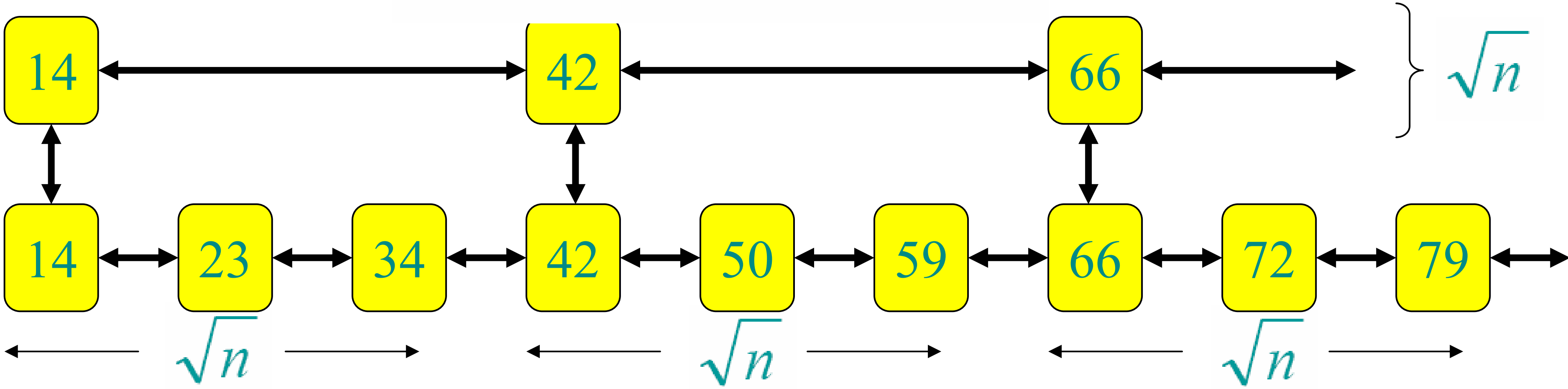


Daha Fazla Bağlantılı Liste

¢ º z¿mleme:

Pekiyi ya, daha fazla sıralı bağlantılı listemiz olsaydı?

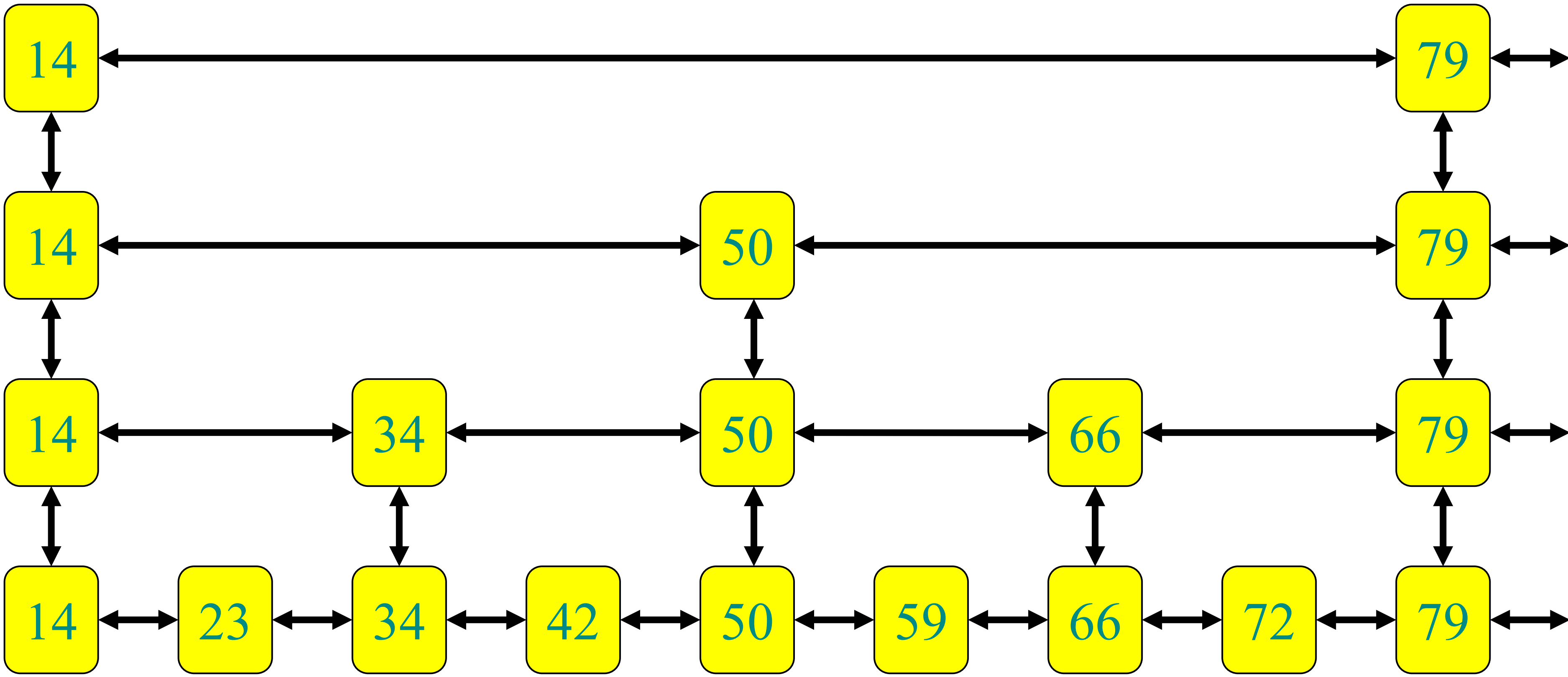
- 2 sıralı liste $\Rightarrow 2 \cdot \sqrt{n}$
- 3 sıralı liste $\Rightarrow 3 \cdot \sqrt[3]{n}$
- k sıralı liste $\Rightarrow k \cdot \sqrt[k]{n}$
- lg n sıralı liste $\Rightarrow \lg n \cdot \sqrt[\lg n]{n} = 2 \lg n$



lg *n* Bağlantılı Liste

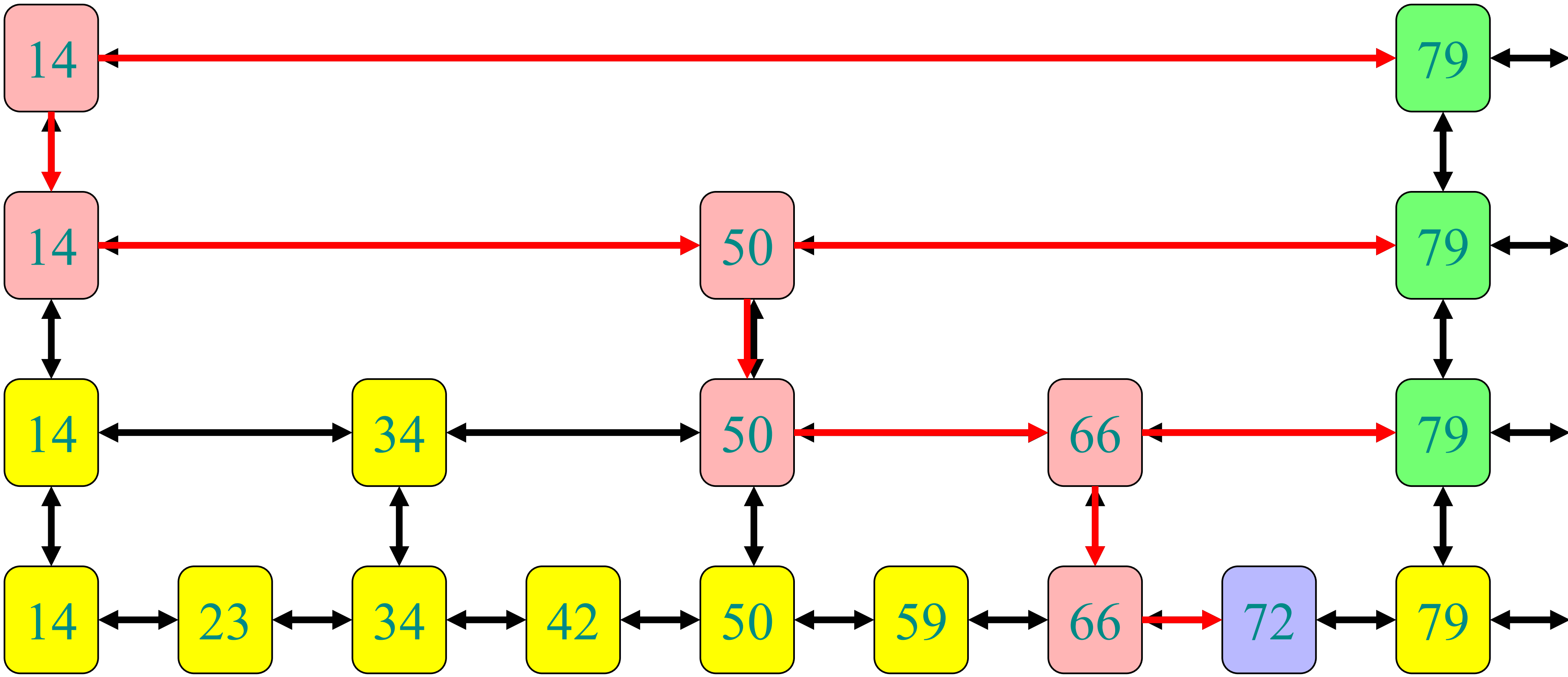
lg *n* sıralı listeler ikili ağaçlar gibidir.

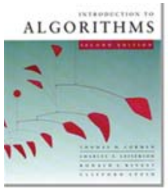
(Aslında seviye bağlantılı B$^+$- ağaçları gibidir.)



lg *n* Bağlantılı Listelerinde Arama

ÖRNEK : ARA (72)

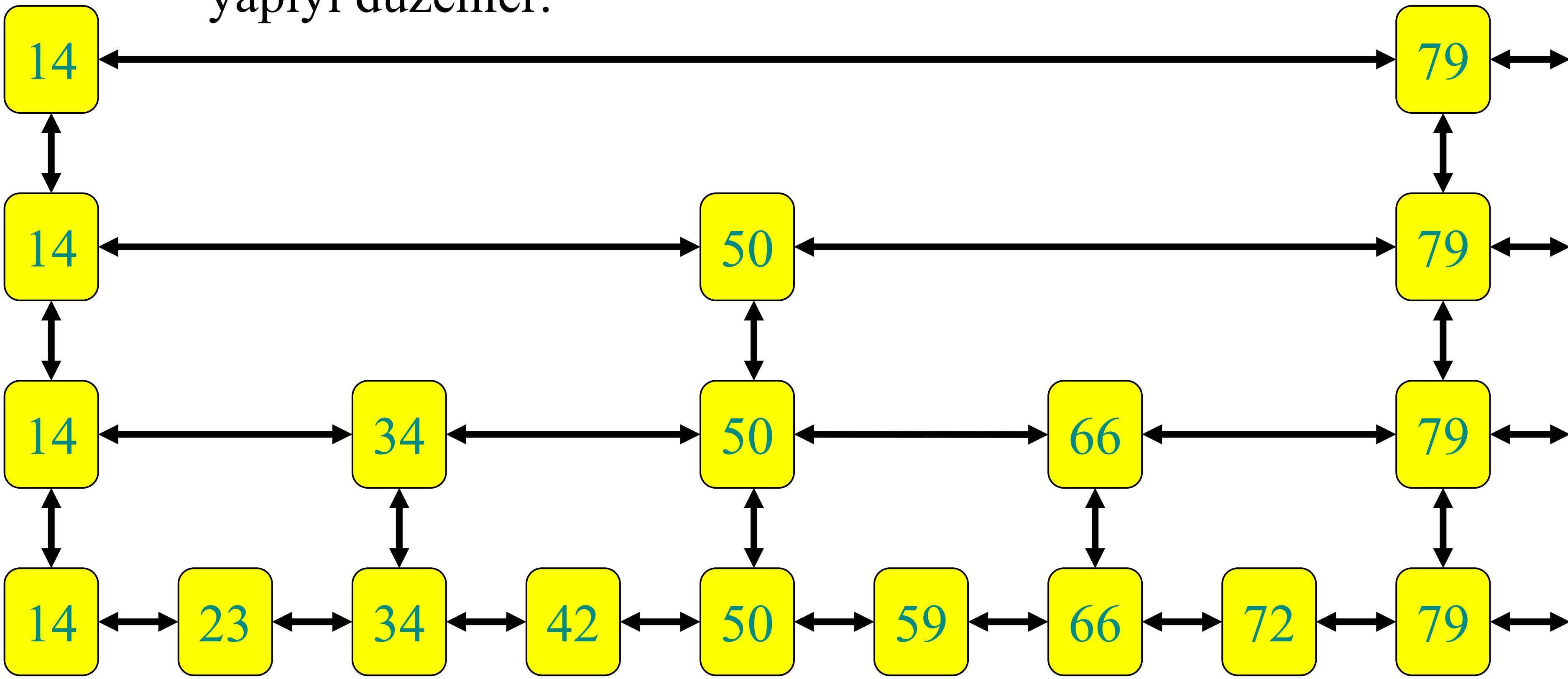


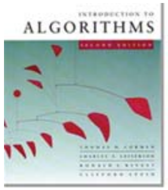


Atlama Listeleri

İdeal Atlama Listesi bu lg *n* yapısındadır.

Atlama listesi veri yapısı, güncellemelerde (ekle/sil) kabaca bu yapıyı düzenler.





ARAYA YERLEŞTİR(x)

Herhangi bir x elemanını Atlama Listesine eklemek için:

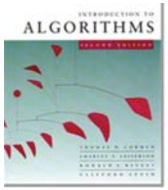
- x' in alt listede nereye denk geldiğini bulmak için ARA(x) yaparız.

- Her zaman en alt listede araya yerleştirme yaparız.

DEĞİŞMEZ : En alt liste bütün elemanları kapsar.

- Üstteki listelerden bazılarına araya yerleştirme yapabiliriz.

SORU : x'i başka hangi listelere eklemeliyiz?



ARAYA YERLEŞTİR (x)

SORU : x'i başka hangi listelerde araya yerleştirmeliyiz?

FİKİR : Yazı-tura atın, eğer YAZI gelirse; x'i bir üst seviyeye yükseltin ve tekrar yazı-tura atın.

- Bir sonraki seviyeye yükselme olasılığı = ½
- Ortalamada;
  - -Elemanların yarısı hiç seviye atlayamaz.
  - -Elemanların 1/4'ü 1 seviye atlar.
  - -Elemanların 1/8'i 2 seviye atlar.
  - -vs.

Yaklaşık olarak dengeli?

Atlama Listesi Örneği

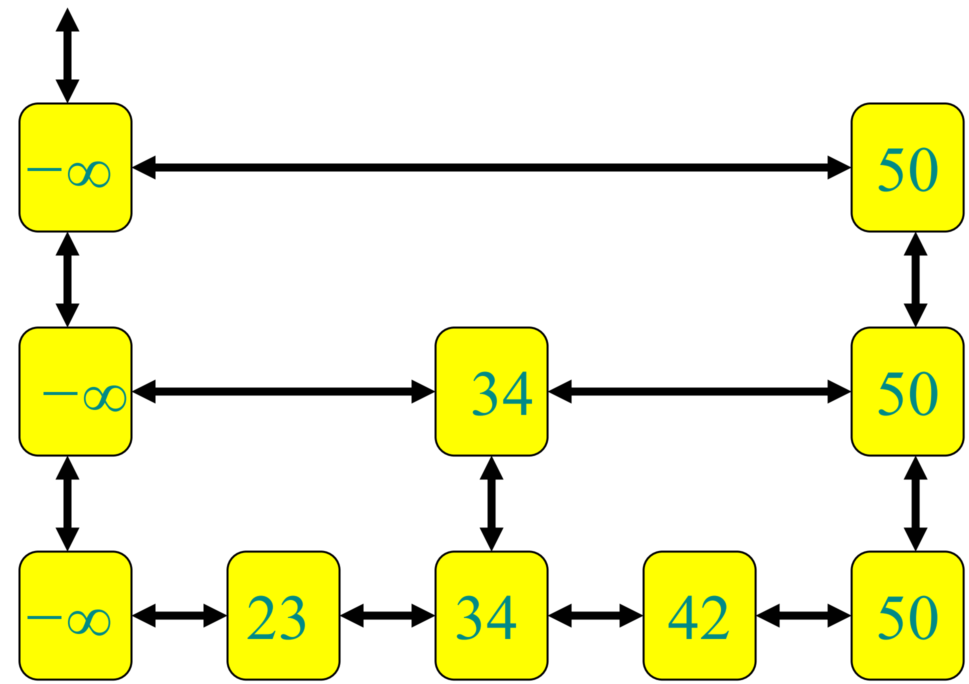
EGZERSİZ : Gerçek bir bozuk para kullanarak, tekrarlanan eklemelerle bir Atlama Listesi yaratın.

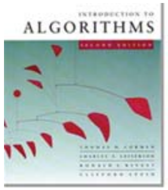
Ufak değişiklik :

• Özel $-\infty$ değerini

her listeye ekleyin.

⇒ Aynı algoritma ile arama yapabilirsiniz.



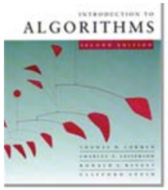


Atlama Listeleri

Boş bir yapıya ($-\infty$'u içeren) yapılan araya yerleştirmeler (ve silmeler) sonucunda oluşan yapı bir ***Atlama Listesi***dir.

- **AR. YER.** (x), rastgele yazı-tura yöntemiyle terfi düzeylerini belirler.

- **SİL**(x), x'in bulunduğu bütün listelerden x'i siler.



Atlama Listeleri

Boş bir yapıya ($-\infty$'u içeren) yapılan araya yerleştirmeler (ve silmeler) sonucunda oluşan yapı bir ***Atlama Listesi***dir.

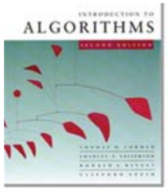
- AR. YER.(x), rastgele yazı-tura yöntemi ile üst seviyeye çıkışları belirler.

- SİL(x), x'in bulunduğu bütün listelerden x'i siler.

Atlama listeleri ne kadar iyidir? (hız/denge)

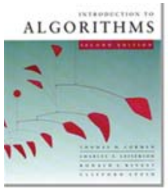
- SEZGİSEL OLARAK : Ortalamada oldukça iyi.

- İDDİA : Hemen her zaman gerçekten, ama gerçekten iyi.



"Yüksek Olasılıkla" Teoremi

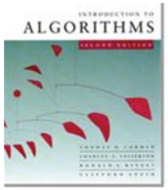
TEOREM : *"Yüksek Olasılıkla"*, n elemanlı bir atlama listesinde her bir arama $O(\lg n)$'ye mal olur.



"Yüksek Olasılıkla" Teoremi

TEOREM : *Yüksek olasılıkla*, n elemanlı bir atlama listesinde her bir arama $O(\lg n)$'e mal olur.

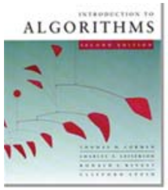
- **Gayri resmi olarak :** Eğer, herhangi bir $\alpha \geq 1$ için, uygun olan ve E olayının $1 - O(1/n^{\alpha})$ ihtimali ile gerçekleştiği bir seçim şansı varsa, E olayı **yüksek olasılıkla (y.o.)** gerçekleşir.
  - Aslında $O(\lg n)$ içindeki sabit, α'ya bağlıdır.
- **Resmi olarak :** Eğer, herhangi bir $\alpha \geq 1$ için, E_{α}'nın en azından $1\text{-}c_{\alpha}/n^{\alpha}$ ihtimali ile gerçekleşmesini sağlayan uygun sabit seçimleri varsa, E_{α} **yüksek olasılıkla** gerçekleşir.



"Yüksek Olasılıkla" Teoremi

TEOREM : *Yüksek olasılıkla*, n elemanlı bir atlama listesinde her bir arama $O(\lg n)$'e mal olur.

- **Gayri resmi olarak :** Eğer, herhangi bir $\alpha \geq 1$ için, uygun olan ve E olayının $1 - O(1/n^{\alpha})$ ihtimali ile gerçekleştiği bir seçim şansı varsa, E olayı **yüksek olasılıkla (y.o.)** gerçekleşir.
- **FİKİR :** α'yı büyük seçerek (örneğin 100), hata olasılığını $O(1/n^{\alpha})$ oldukça azaltabiliriz.
- Hemen hemen kesinlikle, sınırlar, varolan polinomsal zaman algoritmalarında kalacaklardır.



Boole'un Eşitsizliği / Birleşik Sınır

Hatırlatma :

Boole'un Eşitsizliği / Birleşik Sınır

Herhangi bir rastgele olay $E_1, E_2, \ldots, E_k$ için

$$\Pr\{E_1 \cup E_2 \cup \ldots \cup E_k\} \leq \Pr\{E_1\} + \Pr\{E_2\} + \ldots + \Pr\{E_k\}$$

Yüksek olasılıklı olaylarda uygulama:

Eğer $k = n^{O(1)}$ ise ve her E_i olayı yüksek olasılıkla oluşuyorsa, $E_1 \cap E_2 \cap \ldots \cap E_k$

Çözümlemeye Isınma

GERÇEK : *Yüksek Olasılıkla,*

n elemanlı bir atlama listesinin $O(\lg n)$ düzeyi vardır.

KANIT :

- En fazla $c \lg n$ düzeyine sahip olmamızdaki hata ihtimali

$= \Pr\{c \lg n$ düzeyinden daha fazla$\}$.

$\leq n.\ \Pr\{\,x$ elemanı en az $c \lg n$ defa üste çıkmıştır$\}$.

(Boole'un eşitsizliği ile.)

$$= n \cdot (1/2^{c \lg n})$$
$$= n \cdot (1/n^{c})$$
$$= 1/n^{c-1}$$

Çözümlemeye Isınma

GERÇEK : *Yüksek Olasılıkla*,

n elemanlı bir atlama listesinin $O(\lg n)$ düzeyi vardır.

KANIT :

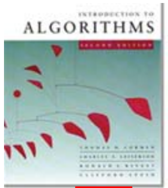
- En fazla $c \lg n$ düzeyine sahip olmamızdaki hata ihtimali

$\leq 1/n^{c-1}$

- Bu olasılık ***polinomsal olarak küçüktür***,

Örnek : en fazla $\alpha = c$-1 için n^{α}

- $O(\lg n)$ sınırındaki sabit c'yi uygun bir şekilde seçerek, α'yı keyfi olarak büyük yapabiliriz.

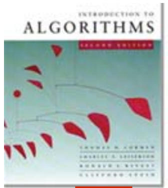


Teoremin İspatı

TEOREM : *Yüksek olasılıkla*, n elemanlı bir atlama listesinde her bir arama $O(\lg n)$'e mal olur.

AKILLICA FİKİR : Arama'yı tersten, yapraktan köke doğru yapmak.

- Arama yaprakta (en alttaki düğüm) başlar. (biter)
- Her düğüm ziyaret edilir:
  - Eğer düğüm bir üste çıkmadıysa (Tura geldiyse), sola gideriz. (soldan gelmiştik)
  - Eğer düğüm bir üste çıktıysa (Yazı geldiyse), yukarı gideriz, (yukarıdan gelmiştik)
- Arama kökte (veya $-\infty$'da) sona erer (başlar).



Teoremin İspatı

TEOREM : *Yüksek olasılıkla*, n elemanlı bir atlama listesinde her bir arama $O(\lg n)$'e mal olur.

AKILLICA FİKİR : Arama'yı tersten, yapraktan köke doğru yapmak.

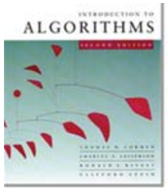
KANIT :

• Arama, köke ulaşana (veya $-\infty$'a) kadar yukarı ve sola ilerler.

•Yukarı hareket sayısı $<$ düzeylerin sayısı

$\leq c \lg n$ y.o. ile (*Önkuram*)

• => y.o. ile, hareket sayısı en fazla $c \lg n$ kere YAZI gelmesi için fırlatmamız gereken para sayısıdır.



Para Atma Analizi

İDDİA: c lg n kere YAZI gelmesi için gereken para atma sayısı

$= \Theta(\lg n)$ y.o. ile

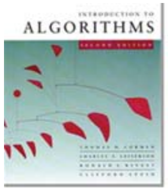
KANIT:

Açıkça $\Omega(\lg n)$: en az $c \lg n$

$O(\lg n)$'i "örnekle" kanıtlayın:

- Diyelim ki; $10\, c \lg n$ atma yaptık.
- En $c \lg n$ kere YAZI ne zaman gelir?

(Daha sonra 10'un rastgele değerlerine genelleyin.)



Para Atma Analizi

İDDİA: c lg n kere YAZI gelmesi için gereken para atma sayısı

$= \Theta(\lg n)$ y.o. ile

KANIT:

- $\Pr\{$tam olarak $c \lg n$ tane YAZI$\} = \underbrace{\binom{10c\lg n}{c\lg n}}_{\text{düzey}} \cdot \underbrace{\left(\frac{1}{2}\right)^{c\lg n}}_{\text{yazı}} \cdot \underbrace{\left(\frac{1}{2}\right)^{9c\lg n}}_{\text{tura}}$

$\Pr\{$ en fazla $c \lg n$ tane YAZI$\} \leq \underbrace{\binom{10c\lg n}{c\lg n}}_{\text{düzeylere aşırı değer biçimi}} \cdot \underbrace{\left(\frac{1}{2}\right)^{9c\lg n}}_{\text{tura}}$

Para Atma Analizi (devam)

Sınırlarla ilgili hatırlatma: $\binom{y}{x}: \quad \left(\frac{y}{x}\right)^x \le \binom{y}{x} \le \left(e\frac{y}{x}\right)^x$

- Pr {en fazla $c \lg n$ YAZI} $\le \binom{10c\lg n}{c\lg n} \cdot \left(\frac{1}{2}\right)^{9c\lg n}$

$$\le \left(e\frac{10c\lg n}{c\lg n}\right)^{c\lg n} \cdot \left(\frac{1}{2}\right)^{9c\lg n}$$

$$= (10e)^{c\lg n} 2^{-9c\lg n}$$

$$= 2^{\lg(10e)\cdot c\lg n} 2^{-9c\lg n}$$

$$= 2^{[\lg(10e)-9]\cdot c\lg n}$$

$$= 1/n^{\alpha} \text{ için } \alpha = [9-\lg(10e)]\cdot c$$

Para Atma Analizi (devam)

- Pr {en fazla $c \lg n$ TURA} $\leq 1/n^{\alpha}$, ($\alpha = [9 - \lg(10e)]c$)

- **ANAHTAR ÖZELLİK:** herhangi bir c için $10 \to \infty$'ki gibi $\alpha \to \infty$

- 10'a ayarlayın, örnek, $O(\lg n)$ sınırında bir sabit, istenen α'ya denk geliyor.

Bu para atma iddiasının ve teoremin kanıtını tamamlıyor.