# CENG 201 Veri Yapıları 8: Özetleme Fonksiyonları(Hash)

Öğr.Gör. Şevket Umut ÇAKIR

Pamukkale Üniversitesi

Hafta 8

#### Anahat

- Bağlı Listesiz Özetleme Doğrusal Sondalama Karesel Sondalama Çift Özetleme Yeniden Özetleme
- Q O(1) Erişime Sahip Özetleme Cuckoo Özetleme Hopscotch Özetleme

#### Özetleme Fonksiyonları

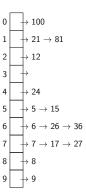
#### Tanım

**Ozetleme(hashing)** işlemi bir veriyi sabit boyutlu bir alana/tabloya izdüşürme işlemidir.

- Örneğin  $f(x) = x \mod 13$  fonksiyonu doğal sayıları 13 uzunluklu bir tabloya izdüşürür
- Özetleme işleminin amacı arama ve ekleme işlemlerinin karmaşıklığını sabit zamana indirmektir

#### Ayrık Zincirleme/Seperate Chaining

21, 7, 8, 9, 17, 6, 24, 26, 27, 81, 12, 36, 5, 15, 100 değerleri eklensin.



#### Doğrusal Sondalama(Deneme)/Linear Probing

- Bağlı listelerde yeni elemanlar için yer ayırma işlemi zaman alıcı olabildiği için ayrık zincirleme dezavantajlı olabilir
- Sırasıyla  $h_0(x), h_1(x), h_2(x), \cdots$  konumlarındaki ilk boş kısma yerleştirilir.  $h_i(x) = (hash(x) + f(i)) \mod TabloBoyutu$
- Özet fonksiyonu bir konum üretir
- Uretilen konum boş ise değer yerleştirilir
- Üretilen konum dolu ise boş bir konum bulunana kadar birer birer sonraki konumlara bakılır
- Doğrusal sondalama için f(i) = i

	Boş tablo		
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			

	Boș tablo	89 sonrası
0		
1		
2		
3		
4		
5		
6		
7		
8		
9		89

# Ekleme Örneği

	Boș tablo	89 sonrası	18 sonrası
0			
1			
2			
3			
4			
5			
6			
7			
8			18
9		89	89

		1		
	Boș tablo	89 sonrası	18 sonrası	49 sonrası
0				49
1				
2				
3				
4				
5				
6				
7				
8			18	18
9		89	89	89

# Ekleme Örneği

Table: Doğrusal Sondalama Ekleme Örneği 89, 18, 49, 58, 69 ekleme

	I				
	Boş tablo	89 sonrası	18 sonrası	49 sonrası	58 sonrası
0				49	49
1					58
2					
3					
4					
5					
6					
7					
8			18	18	18
9		89	89	89	89

Table:  $h(x) = x \mod 10$ 

# Ekleme Örneği

Table: Doğrusal Sondalama Ekleme Orneği 89, 18, 49, 58, 69 ekleme

	Boş tablo	89 sonrası	18 sonrası	49 sonrası	58 sonrası	69 sonrası
0				49	49	49
1					58	58
2						69
3						
4						
5						
6						
7						
8			18	18	18	18
9		89	89	89	89	89

Table:  $h(x) = x \mod 10$ 

# Ekleme Örneği

	Boş tablo	89 sonrası	18 sonrası	49 sonrası	58 sonrası	69 sonrası
0				49	49	49
1					58	58
2						69
3						
4						
5						
6						
7						
8			18	18	18	18
9		89	89	89	89	89

#### Karesel Sondalama(Deneme)/Quadratic Probing

- Sırasıyla  $h_0(x), h_1(x), h_2(x), \cdots$  konumlarındaki ilk boş kısma yerleştirilir.  $h_i(x) = (hash(x) + f(i)) \mod TabloBoyutu$
- Özet fonksiyonu bir konum üretir
- Üretilen konum boş ise değer yerleştirilir
- Uretilen konum dolu ise deneme sayısının karesi kadar ileri bakılır
- Karesel sondalama için  $f(i) = i^2$

```
Boş tablo
0
1
2
3
4
5
6
7
8
9
```

	Boş tablo	89 sonrası
0		
1		
2		
3		
4		
5		
6		
7		
8		
9		89

	Boș tablo	89 sonrası	18 sonrası
0			
1			
2			
3			
4			
5			
6			
7			
8			18
9		89	89

	Boş tablo	89 sonrası	18 sonrası	49 sonrası
0				49
1				
2				
3				
4				
5				
6				
7				
8			18	18
9		89	89	89

Karesel Sondalama

# Ekleme Örneği

Table: Karesel Sondalama Ekleme Örneği 89, 18, 49, 58, 69 ekleme

	Boș tablo	89 sonrası	18 sonrası	49 sonrası	58 sonrası
0	,			49	49
1					
2					58
3					
4					
5					
6					
7					
8			18	18	18
9		89	89	89	89

Table:  $h(x) = x \mod 10$ 

Karesel Sondalama

# Ekleme Örneği

Table: Karesel Sondalama Ekleme Örneği 89, 18, 49, 58, 69 ekleme

	Boș tablo	89 sonrası	18 sonrası	49 sonrası	58 sonrası	69 sonrası
0				49	49	49
1						
2					58	58
3						69
4						
5						
6						
7						
8			18	18	18	18
9		89	89	89	89	89

Table:  $h(x) = x \mod 10$ 

Karesel Sondalama

# Ekleme Örneği

	Boș tablo	89 sonrası	18 sonrası	49 sonrası	58 sonrası	69 sonrası
0				49	49	49
1						
2					58	58
3						69
4						
5						
6						
7						
8			18	18	18	18
9		89	89	89	89	89

Cift Özetleme

# Çift Özetleme/Double Hashing

- Sırasıyla  $h_0(x), h_1(x), h_2(x), \cdots$  konumlarındaki ilk boş kısma yerleştirilir.  $h_i(x) = (hash(x) + f(i)) \mod TabloBoyutu$
- Çakışma durumunda ikinci bir özet fonksiyonu kullanılır
- Çift özetleme için  $f(i) = i \cdot hash_2(x)$
- $hash_2(x) = R (x \mod R)$
- R tablo boyutundan küçük bir asal sayıdır

Cift Özetleme

#### Ekleme Örneği

Table: Çift Özetleme Ekleme Örneği 89, 18, 49, 58, 69 ekleme

```
Boş tablo
0
1
2
3
4
5
6
7
8
9
```

Table: Çift Özetleme Ekleme Örneği 89, 18, 49, 58, 69 ekleme

	Boș tablo	89 sonrası
0		
1		
2		
3		
4		
5		
6		
7		
8		
9		89

Table: Çift Özetleme Ekleme Örneği 89, 18, 49, 58, 69 ekleme

	Boş tablo	89 sonrası	18 sonrası
0			
1			
2			
3			
4			
5			
6			
7			
8			18
9		89	89

Table: Çift Özetleme Ekleme Örneği 89, 18, 49, 58, 69 ekleme

	Boș tablo	89 sonrası	18 sonrası	49 sonrası
0				
1				
2				
3				
4				
5				
6				49
7				
8			18	18
9		89	89	89

Table: Çift Özetleme Ekleme Örneği 89, 18, 49, 58, 69 ekleme

	Boş tablo	89 sonrası	18 sonrası	49 sonrası	58 sonrası
0					
1					
2					
3					58
4					
5					
6				49	49
7					
8			18	18	18
9		89	89	89	89

Table: Çift Özetleme Ekleme Örneği 89, 18, 49, 58, 69 ekleme

	Boș tablo	89 sonrası	18 sonrası	49 sonrası	58 sonrası	69 sonrası
0						69
1						
2						
3					58	58
4						
5						
6				49	49	49
7						
8			18	18	18	18
9		89	89	89	89	89

Table: Çift Özetleme Ekleme Örneği 89, 18, 49, 58, 69 ekleme

	Boș tablo	89 sonrası	18 sonrası	49 sonrası	58 sonrası	69 sonrası
0						69
1						
2						
3					58	58
4						
5						
6				49	49	49
7						
8			18	18	18	18
9		89	89	89	89	89

Cift Özetleme

# Silme İşlemi

 Düğüm yapısında bool isActive alanı tutulur. Bu alan false ise değer silinmiştir

```
public class HashEntry<T> {
1
     T value;
     boolean isActive;
     public HashEntry(T value, boolean isActive) {
4
       this.value = value;
5
       this.isActive = isActive;
6
     public HashEntry(T value) {
       this(value, true);
10
```

Yeniden Özetleme

#### Yeniden Özetleme

- Eğer tablo çok dolarsa ekleme işlemi daha çok zaman alacaktır ve karesel sondalamada uygun konum bulmak imkansız hale gelebilir
- Bu durumda mevcut tablonun en az iki katı büyüklüğünde yeni bir özet tablosu ve özetleme fonksiyonu oluşturulur ve tüm değerler(silinenler hariç) yeni tabloya eklenir

Table: 13, 15, 24, 6, 23 değerleri eklendikten sonra yeni özet tablosu

0	6
1	15
2	23
3	24
4	
5	
6	13

0	
1	
2	
3	
4	
5	
6	6
7	23
8	24
9	
10	
11	
12	
13	13
14	
15	15
16	

Cuckoo Özetleme

#### Cuckoo Özetleme

- Yarısından çoğu boş olan iki adet özet tablosu bulunur
- Her tablo için ayrı bir özetleme fonksiyonu bulunur
- Değerler bu iki tablodan birinde bulunur

Cuckoo Özetleme

# Cuckoo Özetleme Örneği

#### Table: Cuckoo Özetleme Örneği

Tablo 2

Tablo 1				
0	В			
1	С			
2				
3	Е			
4				

0	D
1	
2	Α
3	
4	F

B: 0,0 C: 1,4 D: 1,0 E: 3,2 F: 3,4

A: 0,2

Table: A'nın eklenmesi

Tablo 1				
0	Α			
1				
2				
3				
4				

Tablo 2		
0		
1		
2		
3		
4		

A: 0,2

Table: B'nin eklenmesi

Tablo 1		
0	В	
1		
2		
3		
4		

Tablo 2		
0		
1		
2	Α	
3		
4		

A: 0,2

B: 0,0

Table: C'nin eklenmesi

Tablo 1		
0	В	
1	С	
2		
3		
4		

Tablo 2		
0		
1		
2	Α	
3		
4		

A: 0,2 B: 0,0 C: 1,4 Cuckoo Özetleme

# Ekleme Örneği

Table: D'nin eklenmesi

Tablo 1		
0	В	
1	D	
2		
3		
4		

Tablo 2		
0		
1		
2	Α	
3		
4	С	

Table 2

B: 0,0 C: 1,4 D: 1,0

A: 0,2

Cuckoo Özetleme

# Ekleme Örneği

#### Table: E'nin eklenmesi

Tablo 1		
0	В	
1	D	
2		
3	Е	
4		

Tablo 2		
0		
1		
2	Α	
3		
4	С	

A: 0,2 B: 0,0 C: 1,4 D: 1,0 E: 3,2

Table: F'nin eklenmesi,  $F \rightarrow E$ 

Tablo 1			
0	В		
1	D		
2			
3	F		
4			

Tablo 2		
0		
1		
2	Α	
3		
4	С	

A: 0,2 B: 0,0 C: 1,4 D: 1,0 E: 3,2 F: 3,4

Table: F'nin eklenmesi,  $E \rightarrow A$ 

Tablo 1		
0	В	
1	D	
2		
3	F	
4		

Tablo 2		
0		
1		
2	E	
3		
4	С	

A: 0,2 B: 0,0 C: 1,4 D: 1,0 E: 3,2 F: 3,4

Table: F'nin eklenmesi, A  $\rightarrow$  B

Tablo 1		
0	Α	
1	D	
2		
3	F	
4		

Tablo 2		A: 0,2
0		B: 0,0
1		C: 1,4
2	E	D: 1,0
3		E: 3,2
4	С	F: 3,4

Table: F'nin eklenmesi, B  $\rightarrow$  Tablo2[0]

Tablo 1		
0	Α	
1	D	
2		
3	F	
4		

Tablo 2		A: 0,2
0	В	B: 0,0
1		C: 1,4
2	E	D: 1,0
3		E: 3,2
4	С	F: 3,4

Table: G'nin eklenmesi

0 A 1 D 2 3 F	Tablo 1		
_	0	Α	
_	1	D	
3 F	2		
1	3	F	
4	4		

Tablo 2		
0	В	
1		
2	Е	
3		
4	С	

A: 0,2 B: 0,0 C: 1,4 D: 1,0 E: 3,2 F: 3,4

Table: G'nin eklenmesi

Tablo 1		Tablo 2			
0	Α		0 B		
1	D		1		
2			2	Е	
3	F		3		
4			4	С	

E: 3,2 F: 3,4 G: 1,2

A: 0,2 B: 0,0 C: 1,4 D: 1,0

Table: G  $\rightarrow$  D, D  $\rightarrow$  B, B  $\rightarrow$  A, A  $\rightarrow$  E, E  $\rightarrow$  F, F  $\rightarrow$  C, C  $\rightarrow$  G

Table: G'nin eklenmesi

Tablo 1			
0	В		
1	С		
2			
3	Е		
4			

Tablo 2		
D		
Α		
F		

A: 0,2 B: 0,0 C: 1,4 D: 1,0 E: 3,2 F: 3,4 G: 1,2 Cuckoo Özetleme

### En Fazla Doluluk Oranları

### Table: Çeşitli varyasyonlara göre en fazla doluluk oranları

	Her hücrede 1 eleman	Her hücrede 2 eleman	Her hücrede 4 eleman
2 Özetleme fonksiyonu	0.49	0.86	0.93
3 Özetleme fonksiyonu	0.91	0.97	0.98
4 Özetleme fonksiyonu	0.97	0.99	0.999

### Seksek(Hopscotch) Özetleme

- Klasik doğrusal sondalamayı geliştirmek için ortaya atılmıştır
- Amaç sondalama uzunluğunu kısıtlamaktır
- Bir eleman özgün konumundan en fazla belirli bir uzaklıkta yer alabilir(MAX\_DIST)
- Eğer eleman özgün konumundan çok uzaksa geriye doğru giderek uygun elemanlar yer değiştirilir
- Hop(sıçrama) bilgisi bitlerden oluşur ve hangi uzaklıktaki elemanların mevcut özet değerine sahip olduğunu gösterir
- Orneğin MAX\_DIST=4 için 5. konumdaki 1011 değeri 5,7 ve 8. konumdaki elemanların özet değerinin 5 olduğunu gösterir

Table: Örnek Hopscotch Tablosu

	Eleman	Нор
6	С	1000
7	Α	1100
8	D	0010
9	В	1000
10	Е	0000
11	G	1000
12	F	1000
13		0000
14		0000

A: 7 B: 9 C: 6 D: 7 E: 8 F: 12 G: 11

### Ekleme

- Eklenecek konum boş ise eklenir(hash(x))
- Eklenecek konum boş değilse doğrusal sondalama ile ilk boş konum bulunur:
  - Boş konum özgün konumdan MAX\_DIST uzaklığından az ise eklenir
  - Aksi takdirde boş konum ile MAX\_DIST-1 arasında yer değiştirilebilecek eleman aranır ve istenilen uzaklık değeri elde edilene kadar bu işlem tekrarlanır

Table: H:9 değerinin eklenmesi(MAX\_DIST=4)

	Eleman	Нор
	• • •	
6	С	1000
7	Α	1100
8	D	0010
9	В	1000
10	Е	0000
11	G	1000
12	F	1000
13		0000
14		0000

A: 7

B: 9

C: 6

D: 7

E: 8 F: 12

G: 11

H:9

Table: H:9 değerinin eklenmesi(MAX\_DIST=4)

	Eleman	Нор
6	С	1000
7	А	1100
8	D	0010
9	В	1000
10	Е	0000
11		0010
12	F	1000
13	G	0000
14		0000

A: 7

B: 9

C: 6 D: 7

E: 8

F: 12

G: 11 H:9

Table: H:9 değerinin eklenmesi(MAX\_DIST=4)

	Eleman	Нор
6	С	1000
7	Α	1100
8	D	0010
9	В	1010
10	Е	0000
11	Н	0010
12	F	1000
13	G	0000
14		0000

A: 7

B: 9

C: 6 D: 7

E: 8

F: 12

G: 11 H:9

Table: I:6 değerinin eklenmesi(MAX\_DIST=4)

	Eleman	Нор	
6	С	1000	A: 7
7	Α	1100	B: 9
8	D	0010	C: 6
9	В	1010	D: 7
10	Е	0000	E: 8
11	Н	0010	F: 12
12	F	1000	G: 11
13	G	0000	H:9
14		0000	l:6

Table: I:6 değerinin eklenmesi(MAX\_DIST=4)

	Eleman	Нор	
6	С	1000	A: 7
7	Α	1100	B: 9
8	D	0010	C: 6
9	В	1010	D: 7
10	Е	0000	E: 8
11	Н	0001	F: 12
12	F	1000	G: 1
13		0000	H:9
14	G	0000	l:6

Table: I:6 değerinin eklenmesi(MAX\_DIST=4)

	Eleman	Нор	
	• • •		
6	С	1000	A: 7
7	Α	1100	B: 9
8	D	0010	C: 6
9	В	1010	D: 7
10	Е	0000	E: 8
11	Н	0001	F: 12
12		0100	G: 11
13	F	0000	H:9
14	G	0000	I:6

Table: I:6 değerinin eklenmesi(MAX\_DIST=4)

	Eleman	Нор	
	• • •		
6	С	1000	A: 7
7	Α	1100	B: 9
8	D	0010	C: 6
9		0011	D: 7
10	Е	0000	E: 8
11	Н	0001	F: 12
12	В	0100	G: 11
13	F	0000	H:9
14	G	0000	I:6

Table: I:6 değerinin eklenmesi(MAX\_DIST=4)

	Eleman	Нор	
	• • •		
6	С	1001	A: 7
7	А	1100	B: 9
8	D	0010	C: 6
9	ı	0011	D: 7
10	Е	0000	E: 8
11	Н	0001	F: 12
12	В	0100	G: 11
13	F	0000	H:9
14	G	0000	I:6

