Chapter 17 Amortized Analysis

平摊 (摊还、分摊、分期) 分析

How to analyze algorithms

MERGE-SORT(A, p, r) sorts the elements in the subarray A[p ... r], initially, p = 1, r = n.

ME	ERGE(A, p, q, r)	$T_{\rm sub}(n)$
1	$n_1 \leftarrow q - p + 1$	1
2	$n_2 \leftarrow r - q$	1
3	create arrays $L[1 n_1+1]$ and $R[1 n_2+1]$	1
4	for $i \leftarrow 1$ to n_1	n_1+1
5	$L[i] \leftarrow A[p+i-1]$	n_1
6	for $j \leftarrow 1$ to n_2	n_2+1
7	$R[j] \leftarrow A[q+j]$	n_2
8	$L[n_1+1] \leftarrow \infty$	1
9	$R[n_2+1] \leftarrow \infty$	1
10	$i \leftarrow 1$	1
11	$j \leftarrow 1$	1
12	for $k \leftarrow p$ to r	<i>r-p</i> +2
13	if $L[i] \leq R[j]$	<i>r-p</i> +1
14	$A[k] \leftarrow L[i]$	\boldsymbol{x}
15	$i \leftarrow i+1$	$\boldsymbol{\mathcal{X}}$
16	else $A[k] \leftarrow R[j]$	<i>r</i> - <i>p</i> +1- <i>x</i>
17	$j \leftarrow j+1$	r- p + 1 - x

M	MERGE-SORT(A, p, r)				
1	if $p < r$	1			
2	$q \leftarrow \lfloor (p+r)/2 \rfloor$	1			
3	MERGE-SORT(A, p, q)	T(n/2)			
4	MERGE-SORT(A, q+1, r)	T(n/2)			
5	MERGE(A, p, q, r)	n			

- (1) Sum $T_{\text{sub}}(n) = \sum \dots = n$
- (2) Solving Recursive Equations T(n) = 2T(n/2) + n / master theorem
- (3) Probability Analysis (quick sort)
- (4) ...

An example: Dynamic tables



- We might allocate space for a table, only to find out later that it is not enough. We must then reallocate the table with a larger size and copy all objects stored in the original table over into the new, larger table.

 Expansion
- Similarly, if many objects have been deleted from the table, it may be worthwhile to reallocate the table with a smaller size.



- We assume that the dynamic table supports the operations TABLE-INSERT and TABLE-DELETE.
 - TABLE-INSERT inserts into the table an item that occupies a single slot, that is, a space for one item.
 - Likewise, TABLE-DELETE removes an item from the table, thereby freeing a slot.

Amortized Analysis

- □ In an amortized analysis, we average the time required to perform a sequence of data-structure operations over all the operations performed. 在多个操作中,求一个操作的平均时间
- □ With amortized analysis, we can show that the average cost of an operation is small, if we average over a sequence of operations, even though a single operation within the sequence might be expensive.
- Amortized analysis differs from average-case analysis in that probability is not involved; an amortized analysis guarantees the average performance of each operation in the worst case.
 - 平摊分析不需要考虑输入的概率密度
- □ Amortized cost, 分摊消费: 在一个操作上的平均消费 (cost)

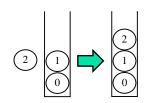
Amortized Analysis

- 17.1 Aggregate analysis (聚集分析)
- 17.2 The accounting method (记账法)
- 17.3 The potential method (势能法)

17.1 Aggregate analysis

• In aggregate analysis, we show that for all n, a sequence of n operations takes worst-case time T(n) in total.

• In the worst case, the average cost, or amortized cost, per operation is therefore T(n)/n.



- PUSH(S, x): pushes object x onto stack S.
- POP(S): pops the top of stack S and returns the popped object. Calling POP on an empty stack generates an error.
- MULTIPOP(*S*, *k*): removes the *k* top objects of stack *S*, popping the entire stack if the stack contains fewer than *k* objects. (The stack has *s* objects.)

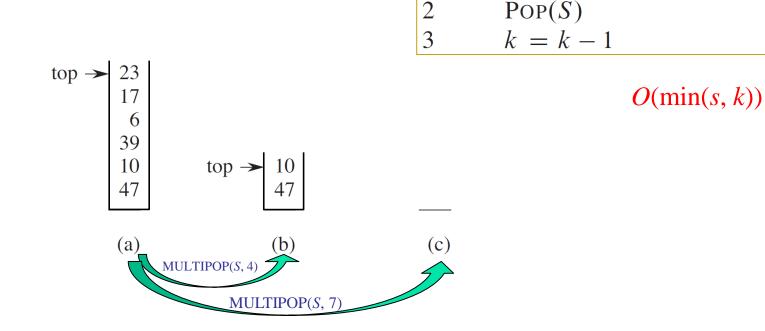
O(1)

 $O(\min(s, k))$?

MULTIPOP(S, k): removes the k top objects of stack S, popping the entire stack if the stack contains fewer than k objects. (The stack has s objects.)

MULTIPOP(S, k)

while not STACK-EMPTY (S) and k > 0



Considering a sequence of *n* PUSH, POP, and MULTIPOP operations on an initially empty stack.

```
STACK(S, n)

1 S = NULL

2 for i \leftarrow 1 to n

3 One of (PUSH(S, i), POP(S), MULIPOP(S, k)
```

```
MULTIPOP(S, k)

1 while not STACK-EMPTY(S) and k > 0

2 POP(S)

3 k = k - 1
```

• Running time? $O(n^2)$

 $O(\min(s, k))$

worst-case: O(n)

- Correct. Not tight.
- Running time tightly? O(n)
- We can pop each object from the stack at most once for each time we have pushed it onto the stack. Therefore, the number of times that POP can be called on a nonempty stack, including calls within MULTIPOP, is at most the number of PUSH operations, which is at most *n*.
 - push一个对象后,至多能被pop一次。但至多n次push。

```
STACK(S, n)

1 S = NULL

2 for i\leftarrow 1 to n

3 One of (PUSH(S, i), POP(S), MULIPOP(S, k)
```

- Running time tightly O(n)
- The average cost of an operation is O(n)/n = O(1)
- In aggregate analysis, the amortized cost of each operation is the average cost.
- For STACK operations, the average cost (the running time) of a stack operation is O(1), we did not use probabilistic reasoning.
- Dividing this total cost by n yielded the average cost per operation, or the amortized cost. 平摊消费(每一个操作的消费)是总消费除以总操作数,即 T(n)/n

Consider the problem of implementing a *k*-bit binary counter that counts upward from 0.







- We use an array A[0...k-1] of bits as the counter. (使用位数组作为计数器)
- A number x that is stored in the counter has its lowest-order bit in A[0] and highest-order

bit in A[k-1], so that $x = \sum_{i=0}^{k-1} A[i] \cdot 2^i$

- Initially, x = 0, and thus A[i] = 0 for all i.
- To add 1 (modulo 2^k) to the value in the counter, we use the following procedures.

```
INCREMENT(A) // 加 1 算法

1 i = 0 // i is the index of A

2 while i < A.length and A[i] == 1

3 A[i] = 0

4 i = i + 1

5 if i < A.length

6 A[i] = 1
```

Counter value	MINGHEN NEW NEW JULY NO.	Total cost
0	$0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0$	0
1	0 0 0 0 0 0 0 1	1
2	0 0 0 0 0 0 1 0	3
3	0 0 0 0 0 0 1 1	4
4	0 0 0 0 0 1 0 0	7
5	0 0 0 0 0 1 0 1	8
6	0 0 0 0 0 1 1 0	10
7	0 0 0 0 0 1 1 1	11
8	0 0 0 0 1 0 0 0	15
9	0 0 0 0 1 0 0 1	16
10	0 0 0 0 1 0 1 0	18
11	0 0 0 0 1 0 1 1	19
12	0 0 0 0 1 1 0 0	22
13	0 0 0 0 1 1 0 1	23
14	0 0 0 0 1 1 1 0	25
15	0 0 0 0 1 1 1 1	26
16	0 0 0 1 0 0 0 0	31

Let us analyze a sequence of n INCREMENT operations on an initially zero counter.
 从 0 开始, 计数到 n

```
COUNTER(A, n)

1 A = 0

2 for j \leftarrow 1 to n

3 INCREMENT(A)
```

```
INCREMENT(A)

1   i = 0

2   while i < A.length and A[i] == 1

3   A[i] = 0

4   i = i + 1

5   if i < A.length

6   A[i] = 1
```

- Running time of COUNTER? O(nk), k是A的位数
- Correct. Not tight.
- Tight running time? O(n)

```
Total
Counter
       MINGKSKAKSKOMINO)
 value
                            cost
        0 0 0 0 0 0 0 0
                             0
        0 0 0 0 0 0 0 1
        0 0 0 0 0 0 1 0
        0 0 0 0 0 0 1 1
        0 0 0 0 0 1 0 0
        0 0 0 0 0 1 0 1
        0 0 0 0 0 1 1 0
                            10
        0 0 0 0 0 1 1 1
                            11
        0 0 0 0 1 0 0 0
                            15
        0 0 0 0 1 0 0 1
                            16
        0 0 0 0 1 0 1 0
                            18
                            19
        0 0 0 0 1 0 1
        0 0 0 0 1 1 0 0
  13
                            23
        0 0 0 0 1 1 0
 14
        0 0 0 0 1 1
                            25
 15
        0 0 0 0 1 1
                            26
        0 0 0 1 0 0 0 0
                            31
```

```
COUNTER(A, n)

1 A = 0

2 for j \leftarrow 1 to n

3 INCREMENT(A)
```

```
INCREMENT(A)

1   i = 0

2   while i < A.length and A[i] == 1

3   A[i] = 0

4   i = i + 1

5   if i < A.length

6   A[i] = 1
```

- Not all bits flip each time INCREMENT is called
- A[0] does flip each time (n times) INCREMENT is called
- A[1], flips only every other time: n/2 times
- A[2], flips only every fourth time: n/4 times
- A[i] flips $n/2^i$ times
- The total number of flips in COUNTER

$$\sum_{i=0}^{k-1} \left\lfloor \frac{n}{2^i} \right\rfloor < n \sum_{i=0}^{\infty} \frac{1}{2^i}$$

$$= 2n$$

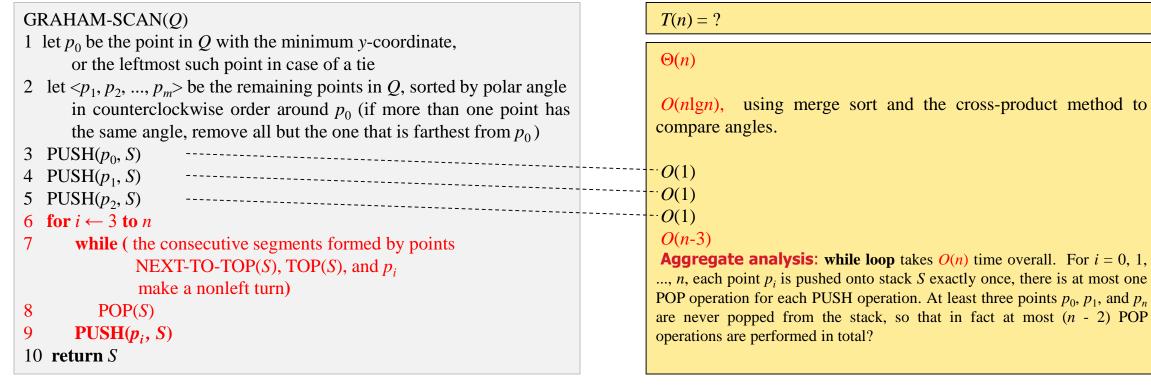
The amortized cost per operation:

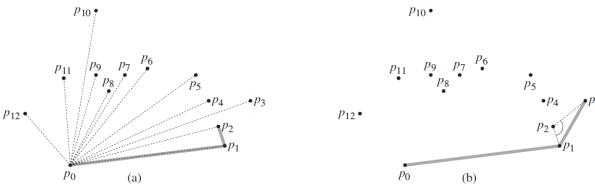
$$O(n)/n = O(1)$$

Counter	A A A A A A A A	Total
value	MINGHSHANSKORING	cost
0	0 0 0 0 0 0 0 0	0
0	0 0 0 0 0 0 0	0
1	$0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1$	1
2	$0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0$	3
3	$0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1$	4
4	$0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0$	7
5	$0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 1$	8
6	$0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 0$	10
7	$0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1$	11
8	0 0 0 0 1 0 0 0	15
9	0 0 0 0 1 0 0 1	16
10	0 0 0 0 1 0 1 0	18
11	$0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 1 \ 1$	19
12	$0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0$	22
13	$0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 0 \ 1$	23
14	0 0 0 0 1 1 1 0	25
15	$0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1$	26
16	0 0 0 1 0 0 0 0	31
	<u>-</u>	

A[1]位, 翻转n/2次; A[0]位, 翻转n次

application of aggregate analysis -- convex hull





17.2 The accounting method

- Amortized cost: the amount we charge an operation.
 - ◆ **Credit**: when an operation's amortized cost exceeds its actual cost, we assign the difference to specific objects in the data structure as credit. 分配的消费券大于实际价格,多余券留下,作为信用,将来用(某些分摊金额小于实际花费时使用)
 - ◆ Credit can help pay for later operations whose amortized cost is less than their actual cost.
- We denote the actual cost of the *i*th operation by c_i and the amortized cost of the *i*th operation by \hat{c}_i , we require

$$\sum_{i=1}^{n} \hat{c}_i \ge \sum_{i=1}^{n} c_i$$

For all sequences of *n* operations.

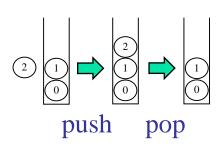
• The total credit

$$\sum_{i=1}^n \widehat{c}_i - \sum_{i=1}^n c_i$$

17.2 The accounting method -(1) Stack operations

- The actual cost of the *i*th operation: c_i
- The amortized cost of the *i*th operation: \hat{c}_i

$$\sum_{i=1}^{n} \widehat{c}_i \ge \sum_{i=1}^{n} c_i$$



• The actual costs of the operations

Pop 1
Multipop min(k, s)

• We assign the amortized costs

Push 2
Pop 0
Multipop 0

Push

• We can pay for any sequence of stack operations by charging the amortized costs?

17.2 The accounting method -(1) Stack operations

- The actual cost of the *i*th operation: c_i
- The amortized cost of the *i*th operation: \hat{c}_i

$$\sum_{i=1}^{n} \widehat{c}_i \ge \sum_{i=1}^{n} c_i$$

• The actual costs of the operations

Pop 1
Multipop min(k, s)

Push

Push

• We assign the amortized costs

Pop

Multipop

- We can pay for any sequence of stack operations.
- For any sequence of *n* Push, Pop, and Multipop operations,

the total amortized cost is O(n), so is the total actual cost.

```
STACK(S, n)

1 S = NULL

2 for i \leftarrow 1 to n

3 One of (PUSH(S, i), POP(S), MULIPOP(S, k)
```

17.2 The accounting method - (2) Incrementing a binary counter

Charge an amortized cost of 2 dollars to set a bit from 0 to 1. When a bit is set, we use 1 dollar (out of the 2 dollars charged) to pay for the actual setting of the bit, and we place the other dollar on the bit as credit to be used later when we flip the bit back to 0. At any point in time, every 1 in the counter has a dollar of credit on it, and thus we can charge nothing to reset a bit to 0.

- ① 0变到1时,支付分摊消费2元(其中1元用于实际转化0到1,结余1元作为信用);
- ② 1变到0时,分摊消费为0元,用信用来支付实际消费(实际翻转)。
- The amortized cost of an INCREMENT is at most 2 dollars.
- For n INCREMENT operations, the total amortized cost is O(n), which bounds the total actual cost.

```
COUNTER(A, n)

1 A = 0

2 for j \leftarrow 1 to n

3 INCREMENT(A)
```

```
INCREMENT(A)

1   i = 0

2   while i < A.length and A[i] == 1

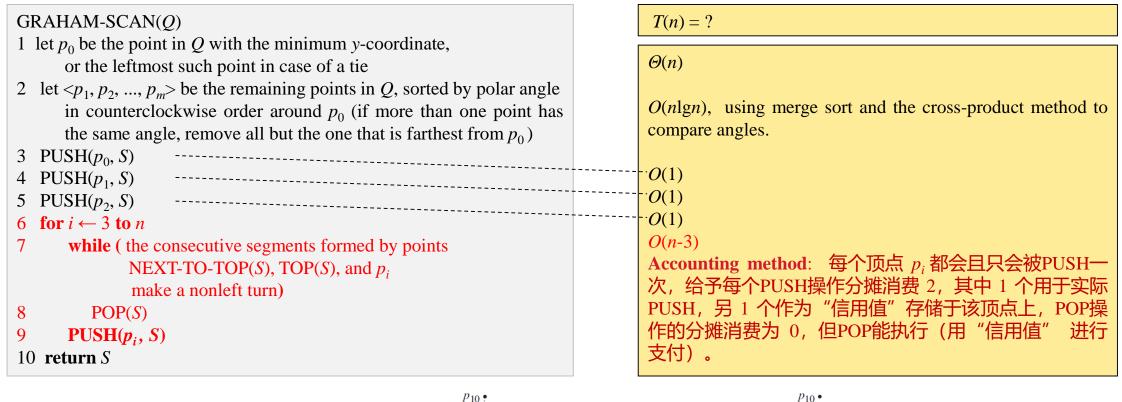
3   A[i] = 0

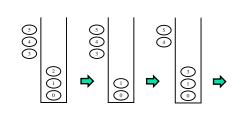
4   i = i + 1

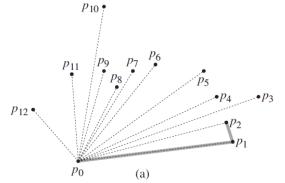
5   if i < A.length

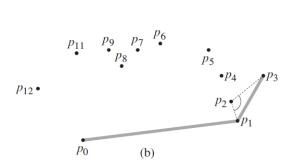
6   A[i] = 1
```

application of accounting method -- convex hull









application of accounting method -- KMP

accounting: q (或 k) 加1时2个分摊消费(1个用于实际消费,1个是信用),减少时0个分摊消费(数q上有q个信用,最多减少到0,因此信用足够支付减少)

```
KMP-MATCHER (T, P)
                                         \Theta(n)
 1 n = T.length
    m = P.length
    \pi = \text{Compute-Prefix-Function}(P)
    q = 0
    for i = 1 to n
        while q > 0 and P[q + 1] \neq T[i]
            q = \pi[q] // cost 0
        if P[q + 1] == T[i]
            q = q + 1 // \cos 2
        if q == m
10
            print "Pattern occurs with shift" i - m
11
12
            q = \pi[q]
```

```
COMPUTE-PREFIX-FUNCTION (P)

1  m = P.length

2  let \pi[1..m] be a new array

3  \pi[1] = 0

4  k = 0

5  for \ q = 2 to \ m

6  while \ k > 0 \ and \ P[k+1] \neq P[q]

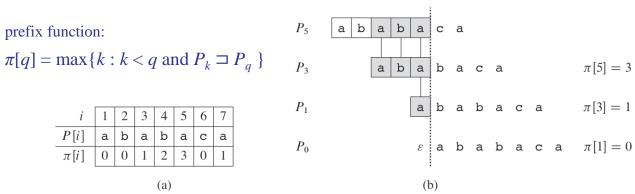
7  k = \pi[k] \ // \cos t \ 0

8  if \ P[k+1] == P[q]

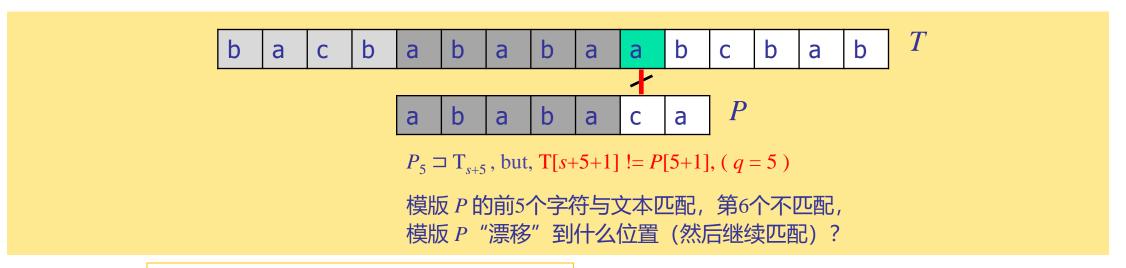
9  k = k+1 \ // \cos t \ 2

10  \pi[q] = k

11  return \ \pi
```



Review of KMP

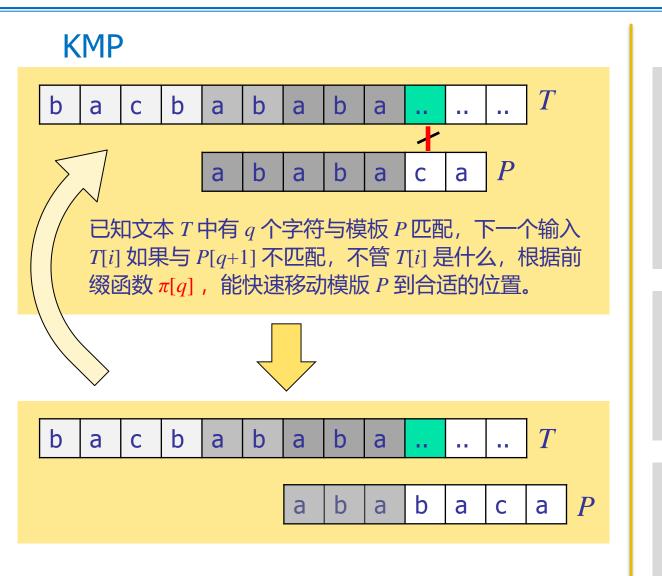


```
KMP-MATCHER(T, P)
 1 \quad n = T.length
 2 m = P.length
 3 \pi = \text{COMPUTE-PREFIX-FUNCTION}(P)
 4 \quad q = 0
   for i = 1 to n
        while q > 0 and P[q + 1] \neq T[i]
            q = \pi[q]
        if P[q + 1] == T[i]
             q = q + 1
10
        if q == m
            print "Pattern occurs with shift" i - m
11
12
             q = \pi[q]
```

Review of KMP



Review: KMP vs FA





When T[i] = b, $\delta(5, b) = 4$, q = 4

Review of KMP

a

a

a

C

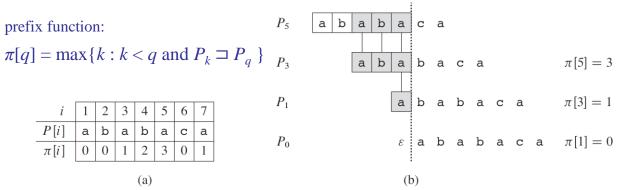
a

如何求 $\pi[q]$? KMP-MATCHER, COMPUTE-PREFIX-FUNCTION, 这两个函数的逻辑完全相同,第一个是找 T_x 最大后缀,第二个找 P_a 的最大后缀(两者都是 P 的前缀)。

```
KMP-MATCHER (T, P)
                                          \Theta(n)
 1 \quad n = T.length
    m = P.length
    \pi = \text{Compute-Prefix-Function}(P)
    q = 0
    for i = 1 to n
        while q > 0 and P[q + 1] \neq T[i]
            q = \pi[q] // cost 0
        if P[q + 1] == T[i]
            q = q + 1 // \cos 2
        if q == m
10
             print "Pattern occurs with shift" i - m
11
12
             q = \pi[q]
                                                                 prefix function:
a
         b
              a
                   b
                        a
                                  a
```

```
COMPUTE-PREFIX-FUNCTION(P)

1 \quad m = P.length
2 \quad let \pi[1 .. m] \text{ be a new array}
3 \quad \pi[1] = 0
4 \quad k = 0
5 \quad \text{for } q = 2 \text{ to } m
6 \quad \text{while } k > 0 \text{ and } P[k+1] \neq P[q]
7 \quad k = \pi[k] \text{ // cost } 0
8 \quad \text{if } P[k+1] == P[q]
9 \quad k = k+1 \text{ // cost } 2
10 \quad \pi[q] = k
11 \quad \text{return } \pi
```



17.3 The potential method (势能法)

- The potential method of amortized analysis represents the prepaid work as "potential energy" (potential), which can be released to pay for future operations.
- We associate the potential with the data structure (DS) as a whole rather than with specific objects within the data structure.
- The potential method works as follows:
 - We will perform n operations, starting with an initial DS D_0 .
 - c_i : the actual cost of the *i*th operation (i = 1, 2, ..., n).
 - D_i : the DS that results after applying the *i*th operation to DS D_{i-1} .
 - Φ : A potential function Φ maps D_i to a real number $\Phi(D_i)$, which is the potential associated with D_i .
- The amortized cost \hat{c}_i of the *i*th operation with respect to potential function Φ is defined by

$$\hat{c}_i = c_i + \Phi(D_i) - \Phi(D_{i-1})$$







17.3 The potential method (势能法)

- The potential method of amortized analysis represents the prepaid work as "potential energy" (potential), which can be released to pay for future operations.
- We associate the potential with the data structure (DS) as a whole rather than with specific objects within the data structure.
- The potential method works as follows:
 - We will perform n operations, starting with an initial DS D_0 .
 - c_i : the actual cost of the *i*th operation (i = 1, 2, ..., n).
 - D_i : the DS that results after applying the *i*th operation to DS D_{i-1} .
 - Φ : A potential function Φ maps D_i to a real number $\Phi(D_i)$, which is the potential associated with D_i .
- The amortized cost \hat{c}_i of the *i*th operation with respect to potential function Φ is defined by

$$\hat{c}_i = c_i + \Phi(D_i) - \Phi(D_{i-1})$$

• The total amortized cost of the *n* operations is

$$\sum_{i=1}^{n} \hat{c}_{i} = \sum_{i=1}^{n} (c_{i} + \Phi(D_{i}) - \Phi(D_{i-1}))$$
$$= \sum_{i=1}^{n} c_{i} + \Phi(D_{n}) - \Phi(D_{0})$$

17.3 The potential method (势能法)

- The amortized cost \hat{c}_i of the *i*th operation with respect to potential function Φ is defined by $\hat{c}_i = c_i + \Phi(D_i) \Phi(D_{i-1})$
- The total amortized cost of the *n* operations is

$$\sum_{i=1}^{n} \hat{c}_{i} = \sum_{i=1}^{n} (c_{i} + \Phi(D_{i}) - \Phi(D_{i-1}))$$
$$= \sum_{i=1}^{n} c_{i} + \Phi(D_{n}) - \Phi(D_{0})$$

- We usually just define $\Phi(D_0)$ to be 0 and then show that $\Phi(D_i) \ge 0$ for all i.
- Different potential functions may yield different amortized costs.

17.3 The potential method - (1) Stack operations

• The amortized cost \hat{c}_i of the *i*th operation with respect to potential function Φ is defined by

$$\hat{c}_i = c_i + \Phi(D_i) - \Phi(D_{i-1})$$

• The total amortized cost of the *n* operations is $\sum_{i=1}^{n} \hat{c}_{i} = \sum_{i=1}^{n} (c_{i} + \Phi(D_{i}) - \Phi(D_{i-1}))$ $= \sum_{i=1}^{n} c_{i} + \Phi(D_{n}) - \Phi(D_{0})$

```
STACK(S, n)

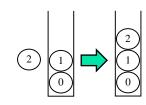
1 S = NULL

2 for i \leftarrow 1 to n

3 One of (PUSH(S, i), POP(S), MULIPOP(S, k))
```

- We define the potential function Φ on a stack to be the number of objects in the stack.
 - for the empty stack D_0 with which we start, we have $\Phi(D_0) = 0$
 - after the *i*th operation, for the stack D_i , $\Phi(D_i) \ge 0$?
- If the *i*th operation on a stack containing *s* objects is a PUSH operation, the potential difference is $\Phi(D_i)$ $\Phi(D_{i-1}) = (s+1)$ s=1.
- The amortized cost of this **PUSH** operation is

$$\hat{c}_i = c_i + \Phi(D_i) - \Phi(D_{i-1}) = 1 + 1 = 2$$



17.3 The potential method - (1) Stack operations







- If the *i*th operation is **MULTIPOP**(S, k), which causes $k' = \min(k, s)$ objects to be popped off the stack. The actual cost of the operation is k', and the potential difference is $\Phi(D_i) \Phi(D_{i-1}) = -k'$
- The amortized cost of this MULTIPOP operation is

$$\hat{c}_i = c_i + \Phi(D_i) - \Phi(D_{i-1}) = k' - k' = 0$$

- Similarly, the amortized cost of an ordinary **POP** operation is 0.
- The worst-case cost of
 STACK is therefore O(n)?

```
STACK(S, n)

1 S = NULL

2 for i \leftarrow 1 to n

3 One of (PUSH(S, i), POP(S), MULIPOP(S, k)) // 2 at most
```

17.3 The potential method - (2) Incrementing a binary counter

- We define the potential Φ of the counter after the *i*th INCREMENT operation to be b_i , the number of 1s in the counter. 第 i 次 "INCREMENT 加1" 操作后,势函数 $b_i = \Phi(D_i)$ 为计数器中 1 的个数
- Suppose that the *i*th INCREMENT operation **resets** t_i bits (**from 1 to 0**). The actual cost of the operation is therefore at most, that is $c_i \le t_i + 1$, since in addition to resetting t_i bits, it sets at most one bit to 1. 实际的消费 $c_i \le t_i + 1$, 即,把 $t_i \land 1$ 变为 0,至多把 1 个 0 变为 1
 - If $b_i = 0$, then the *i*th operation resets all *k* bits. 计数器中 1 的个数为0,即第 *i* 次操作重置所有 *k* 位。原 *k* 位都是 1,已经是最大数,加 1 操作后越界限,变成 1000..00,最高位 1 舍去,因此 $b_{i-1} = t_i = k$. (第 *i* 次操作前有 *k* 个 1)
 - If $b_i > 0$, then $b_i = b_{i-1} t_i + 1$. 有效计数范围内, t_i 个1变为0 (while循环,即第2~4行),1个0变为1 (加1算法的最后一行,即第6行)
 - In either case, $b_i \le b_{i-1} t_i + 1$, and the potential difference is

$$\Phi(D_i) - \Phi(D_{i-1}) = b_i - b_{i-1} \le (b_{i-1} - t_i + 1) - b_{i-1} = 1 - t_i$$

$$\hat{c}_i = c_i + \Phi(D_i) - \Phi(D_{i-1}) \le (t_i + 1) + (1 - t_i) = 2$$

• The amortized cost of **INCREMENT** is therefore 2.

```
COUNTER(A, n)

1 A = 0

2 for j \leftarrow 1 to n

3 INCREMENT(A)
```

```
INCREMENT(A)

1  i = 0

2  while i < A.length and A[i] == 1

3  A[i] = 0

4  i = i + 1

5  if i < A.length

6  A[i] = 1
```

```
        Counter value
        Total cost

        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        1
        1
        2
        0
        0
        0
        0
        0
        0
        0
        0
        1
        1
        2
        0
        0
        0
        0
        0
        1
        1
        4
        4
        0
        0
        0
        0
        0
        1
        1
        4
        4
        0
        0
        0
        0
        1
        0
        0
        7
        5
        0
        0
        0
        0
        1
        0
        0
        0
        1
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
        0
```

17.3 The potential method - (2) Incrementing a binary counter

- We define the potential Φ of the counter after the *i*th INCREMENT operation to be b_i , the number of 1s in the counter. 第 i 次操作后,势函数 b_i 为计数器中 1 的个数
- The amortized cost is therefore

$$\hat{c}_i = c_i + \Phi(D_i) - \Phi(D_{i-1}) \le (t_i + 1) + (1 - t_i) = 2$$

```
COUNTER(A, n)

1 A = 0

2 for j \leftarrow 1 to n

3 INCREMENT(A)
```

```
INCREMENT(A)

1   i = 0

2   while i < A.length and A[i] == 1

3   A[i] = 0

4   i = i + 1

5   if i < A.length

6   A[i] = 1
```

• The worst-case cost of COUNTER is therefore O(n)?

17.4 Dynamic tables



- We might allocate space for a table, only to find out later that it is not enough. We must then reallocate the table with a larger size and copy all objects stored in the original table over into the new, larger table.

 Expansion
- Similarly, if many objects have been deleted from the table, it may be worthwhile to reallocate the table with a smaller size.



- We assume that the dynamic table supports the operations TABLE-INSERT and TABLE-DELETE.
 - TABLE-INSERT inserts into the table an item that occupies a single slot, that is, a space for one item.
 - Likewise, TABLE-DELETE removes an item from the table, thereby freeing a slot.

17.4 Dynamic tables

• The details of the data-structuring method used to organize the table are unimportant. It might be:

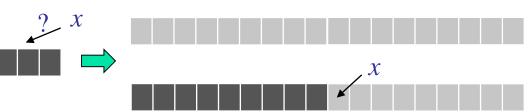
a stack, or a heap, or a hash table, or an array 数据结构用什么没关系,可以是栈、堆、hash表、数组、...

• load factor $\alpha(T)$ of a nonempty table T: $\alpha(T) = T.num/T.size$

• We assign an empty table (one with no items) size 0, and we define its load factor to be 1.

17.4.1 Table expansion

- We assume that storage for a table is allocated as an **array** of slots.
- Upon inserting an item into a full table, we can expand the table by allocating a new table with more slots than the old table had. 往一个满表插入新元素时,需要先进行表扩展
- A common heuristic allocates a new table with twice as many slots as the old one.
- If the only table operations are **insertions**, then the load factor of the table is always at least 1/2, and thus the amount of wasted space never exceeds half the total space in the table.



往一个满表插入新元素时,需要先进行表扩展(一般是容量倍增),然后把原来的元素 拷贝到扩展后的表,然后再插入新元素。

```
TABLE-INSERT (T, x)
     if T. size == 0
          allocate T.table with 1 slot
          T.size = 1
     if T.num == T.size
          allocate new-table with 2 \cdot T. size slots
         insert all items in T.table into new-table
 6
         free T. table
          T.table = new-table
          T.size = 2 \cdot T.size
 9
     insert x into T.table
10
                                  < T.table: a pointer to T >
11
     T.num = T.num + 1
```

17.4.1 Table expansion

从 0 开始, 插入 *n* 个元素, The running time of TABLE-EXPANSION?

```
TABLE-EXPANSION(T, n)

1 T = NULL

2 for i \leftarrow 1 to n

3 TABLE-INSERT(T, x)
```

```
TABLE-INSERT (T, x)

1 if T.size == 0

2 allocate T.table with 1 slot

3 T.size = 1

4 if T.num == T.size

5 allocate new-table with 2 \cdot T.size slots

6 insert all items in T.table into new-table

7 free T.table

8 T.table = new-table

9 T.size = 2 \cdot T.size

10 insert x into T.table

11 T.num = T.num + 1
```



- What is the cost c_i of the *i*th operation?
 - If the current table has room for the new item (or the first operation), then $c_i = 1$?
 - If the current table is full, an expansion occurs, then $c_i = i$?
- The worst-case cost of an operation is O(n)?
- The upper bound of TABLE-EXPANSION is $O(n^2)$?
- Correct. Not tight?

17.4.1 Table expansion – (1) aggregate analysis

从 0 开始,插入 *n* 个元素,The running time of TABLE-EXPANSION?

```
TABLE-EXPANSION(T, n)

1 T = NULL

2 for i \leftarrow 1 to n

3 TABLE-INSERT(T, x)
```

```
TABLE-INSERT (T, x)

1 if T.size == 0

2 allocate T.table with 1 slot

3 T.size = 1

4 if T.num == T.size

5 allocate new-table with 2 \cdot T.size slots

6 insert all items in T.table into new-table

7 free T.table

8 T.table = new-table

9 T.size = 2 \cdot T.size

10 insert x into T.table

11 T.num = T.num + 1
```

What is the cost c_i of the *i*th operation?

$$c_{i} = \begin{cases} i, & \text{if } i\text{-}1=2^{k}, \\ 1, & \text{otherwise.} \end{cases} \Rightarrow \sum_{i=1}^{n} c_{i} \leq n + \sum_{j=0}^{\lfloor \lg n \rfloor} 2^{j} \\ < n+2n \\ = 3n \end{cases}$$

No.	cost	T							
1	1								
2	2								
3	3								
4	1								
5	5								
6	1								

Red: 执行插入元素操作(次数)

Gray: 空槽 null slot

Black: 有元素的槽

17.4.1 Table expansion – (1) aggregate analysis

The running time of TABLE-EXPANSION?

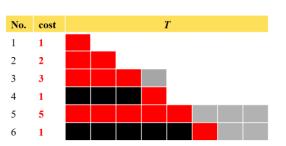
```
TABLE-EXPANSION(T, n)

1 T = NULL

2 for i \leftarrow 1 to n

3 TABLE-INSERT(T, x)
```

TABLE-INSERT (T, x)1 if T.size == 02 allocate T.table with 1 slot 3 T.size = 14 if T.num == T.size5 allocate new-table with $2 \cdot T.size$ slots 6 insert all items in T.table into new-table7 free T.table8 T.table = new-table9 $T.size = 2 \cdot T.size$ 10 insert x into T.table11 T.num = T.num + 1



What is the cost c_i of the *i*th operation?



17.4.1 Table expansion – (2) accounting method

从 0 开始, 插入 *n* 个元素, The running time of TABLE-EXPANSION?

```
TABLE-EXPANSION(T, n)

1 T = NULL

2 for i \leftarrow 1 to n

3 TABLE-INSERT(T, x)
```

```
TABLE-INSERT (T, x)

1 if T.size == 0

2 allocate T.table with 1 slot

3 T.size = 1

4 if T.num == T.size

5 allocate new-table with 2 \cdot T.size slots

6 insert all items in T.table into new-table

7 free T.table

8 T.table = new-table

9 T.size = 2 \cdot T.size

10 insert x into T.table

11 T.num = T.num + 1
```

What is the cost c_i of the *i*th operation?

$$c_{i} = \begin{cases} i, & \text{if } i\text{-}1=2^{k}, \\ 1, & \text{otherwise.} \end{cases} \implies \sum_{i=1}^{n} c_{i} \leq n + \sum_{j=0}^{\lfloor \lg n \rfloor} 2^{j}$$

$$< n + 2n$$

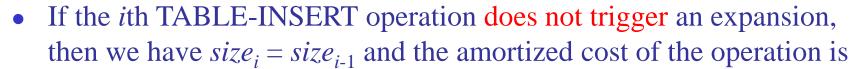
$$= 3n$$

We charge an amortized cost of 3 dollars for each insertion, 1 for actual insertion, 2 for credits.

No.	cost	T							
1	1	2	total is 3: 1 for cost, 2 left for credits.						
2	2	1	2						
3	3	0	1	2					
4	1	0	1	2	2				
5	5	-1	0	1	1	2			
6	1	-1	0	1	1	2	2		

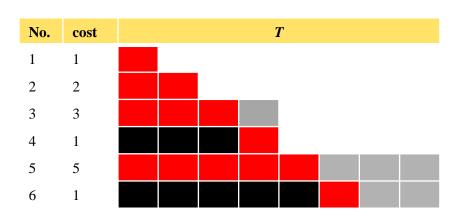
17.4.1 Table expansion – (3) potential method

- potential function : $\Phi(T) = 2T.num T.size$
 - Immediately after an expansion, we have T.num = T.size/2, thus $\Phi(T) = 0$.
 - Immediately before an expansion, we have T.num = T.size, thus $\Phi(T) = T.num$.



$$\hat{c}_{i} = c_{i} + \Phi_{i} - \Phi_{i-1}
= 1 + (2 \cdot num_{i} - size_{i}) - (2 \cdot num_{i-1} - size_{i-1})
= 1 + (2 \cdot num_{i} - size_{i}) - (2(num_{i} - 1) - size_{i})
= 3$$

TAF	TABLE-EXPANSION(T , n)					
1 <i>T</i>	= NULL					
2 f c	or $i \leftarrow 1$ to n					
3	TABLE-INSERT(T , x)					



17.4.1 Table expansion – (3) potential method

- potential function : $\Phi(T) = 2T.num T.size$
 - Immediately after an expansion, we have T.num = T.size/2, thus $\Phi(T) = 0$.
 - Immediately before an expansion, we have T.num = T.size, thus $\Phi(T) = T.num$.



$$size_i = 2size_{i-1}$$
, and $size_{i-1} = num_{i-1} = num_i - 1$, thus

$$\hat{c}_{i} = c_{i} + \Phi_{i} - \Phi_{i-1}
= num_{i} + (2 \cdot num_{i} - size_{i}) - (2 \cdot num_{i-1} - size_{i-1})
= num_{i} + (2 \cdot num_{i} - 2 \cdot (num_{i} - 1)) - (2(num_{i} - 1) - (num_{i} - 1))
= num_{i} + 2 - (num_{i} - 1)
- 3$$

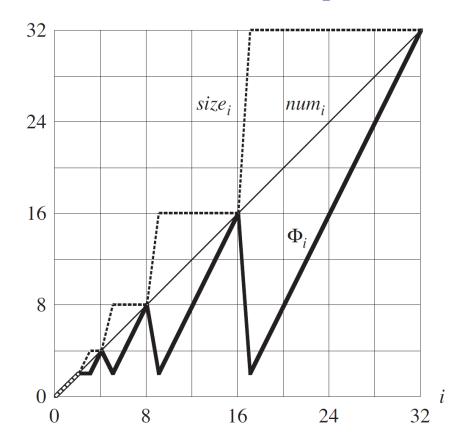
TABLE-EXPANSION(T , n)			
1 T = NULL			
2 for $i \leftarrow 1$ to n			
3 TABLE-INSERT(T , x)			

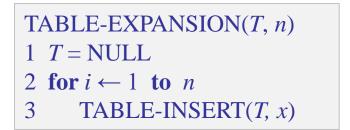
No.	cost		7	Γ		
1	1					
2	2					
3	3					
4	1					
5	5					
6	1					

17.4.1 Table expansion – (3) potential method

- potential function : $\Phi(T) = 2T.num T.size$
 - Immediately after an expansion, we have T.num = T.size/2, thus $\Phi(T) = 0$.
 - Immediately before an expansion, we have T.num = T.size, thus $\Phi(T) = T.num$.







No.	cost	T					
1	1						
2	2						
3	3						
4	1						
5	5						
6	1						

17.4.2 Table contraction (表缩小)



exercises

- 所有的课后习题。
- 再举几个例子(找凸包、KMP, 我已经用过了,不能再用),说明如何用Amortized Analysis进行算法的复杂度分析。
- 把本书中所有用Amortized Analysis进行复杂度分析的算法找出来,并指出其分析方法与结果。

exercises

把本书(也可以不限于本书)中所有用Amortized Analysis进行复杂度分析是算法找出来,并指出其分析方法与结果(可作为期末讲解的大作业)。

