DATA STRUCTURES AND ALGORITHMS LECTURE 7

Lect. PhD. Marian Zsuzsanna

Babeş - Bolyai University Computer Science and Mathematics Faculty

2017 - 2018



In Lecture 6...

- Linked Lists on Arrays
- Skip Lists
- ADT Set
- ADT Map
- Iterator

Today

- ADT Matrix
- 2 ADT List

3 ADT Stack

ADT Matrix

- A *Matrix* is a container that represents a two-dimensional array.
- Each element has a unique position, determined by two indexes: its line and column.

ADT Matrix

- A Matrix is a container that represents a two-dimensional array.
- Each element has a unique position, determined by two indexes: its line and column.
- The operations for a Matrix are different from the operations that exist for most other containers, because in a Matrix we cannot add elements, and we cannot delete an element from a Matrix, we can only change the value of an element.

Matrix - Operations

- The minimum set of operations that should exist for the ADT Matrix is:
 - init(matrix, nrL, nrC) create a new matrix with nrL lines and nrC columns
 - nrLine(matrix) return the number of lines from the matrix
 - nrColumns(matrix) return the number of columns from the matrix
 - element(matrix, i, j) return the element from the line i and column j
 - modify(matrix, i, j, val) change the values of the element from line i and column j into val

Matrix - Operations

- Other possible operations:
 - get the position of a given element
 - create an iterator that goes through the elements by columns
 - create an iterator the goes through the elements by lines
 - etc.

Matrix - representation

- Usually a sequential representation is used for a Matrix (we memorize all the lines one after the other in a consecutive memory block).
- If the Matrix contains many values of 0 (or 0_{TElem}), we have a sparse matrix, where it is more (space) efficient to memorize only the elements that are different from 0.

Sparse Matrix example

• Out of the 36 elements, only 10 are different from 0.

Sparse Matrix - representation

- We can memorize (line, column, value) triples, where value is different from 0 (or 0_{TElem}). For efficiency, we memorize the elements sorted by the (line, column) pairs (if the lines are different we order by line, if they are equal we order by column).
- Triples can be stored in:
 - (dynamic) arrays
 - linked lists
 - (balanced) binary trees

- For the previous example we would keep the following triples: <1,3,3>, <1,5,5>, <2,1,2>, <3,6,4>, <4,1,1>, <4,4,7>, <5,2,6>, <5,6,5>, <6,3,9>, <6,4,1>,
- We need to retain the dimensions of the matrix as well (we might have last line(s) or column(s) with only 0 values).

Sparse Matrix - representation

- Linked representation, using circular lists.
- Each node contains the line, the column, and the value (different from 0) and each node has two pointers: to the next element on the same line and to the next element on the same column. Last elements keep a pointer to the first ones (circular lists).
- We will have special nodes for each line and each column to show the beginning of the corresponding list.

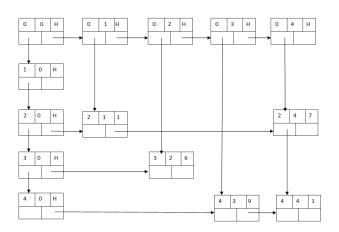
For the following matrix:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 7 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 9 & 1 \end{bmatrix}$$

 The linked lists will be made of nodes. Each node contains the line, column and value and two pointers: one to the next element on the same line, and one to the next element from the same column.

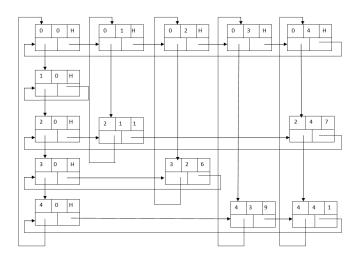


• This is how a node will be represented on the following figures. It represents the element from line 2, column 1 with value 1.



- Nodes with line or column 0 and with value H, are header nodes, they do not represent actual elements, just the first node of the corresponding column or row.
- Obviously, the nodes can be anywhere in the memory (but it is easier to understand the representation if we draw them like this).
- And since we have circular lists, on each row and column the last node has a pointer to the corresponding header node.
- It is enough to retain the address of header node 0,0.





Sparse Matrix - operations

 Operations of a sparse matrix are exactly the same as the operations for a regular matrix. The most difficult operation is modify, because here we have 4 different cases, based on the current value at line i and column j (we will call it old_value) and the value we want to put there (new_value).

Sparse Matrix - operations

- Operations of a sparse matrix are exactly the same as the operations for a regular matrix. The most difficult operation is modify, because here we have 4 different cases, based on the current value at line i and column j (we will call it old_value) and the value we want to put there (new_value).
 - $old_value = 0$ and $new_value = 0 \Rightarrow$ do nothing
 - old_value = 0 and new_value ≠ 0 ⇒ add a new triple/node with new_value
 - $old_value \neq 0$ and $new_value = 0 \Rightarrow$ delete the triple/node with old_value
 - old_value ≠ 0 and new_value ≠ 0 ⇒ modify the value from the triple/node to new_value



ADT List

- A *list* can be seen as a sequence of elements of the same type, $\langle l_1, l_2, ..., l_n \rangle$, where there is an order of the elements, and each element has a *position* inside the list.
- In a list, the order of the elements is important (positions are important).
- The number of elements from a list is called the length of the list. A list without elements is called *empty*.

ADT List

- A List is a container which is either empty or
 - it has a unique first element
 - it has a unique last element
 - for every element (except for the last) there is a unique successor element
 - for every element (except for the first) there is a unique predecessor element
- In a list, we can insert elements (using positions), remove elements (using positions), we can access the successor and predecessor of an element from a given position, we can access an element from a position.



ADT List - Positions

- Every element from a list has a unique position in the list:
 - positions are relative to the list (but important for the list)
 - the position of an element:
 - identifies the element from the list
 - determines the position of the successor and predecessor element (if they exist).

ADT List - Positions

- Position of an element can be seen in different ways:
 - as the rank of the element in the list (first, second, third, etc.)
 - similarly to an array, the position of an element is actually its index
 - as a reference to the memory location where the element is stored.
 - for example a pointer to the memory location
- For a general treatment, we will consider in the following the position of an element in an abstract manner, and we will consider that positions are of type TPosition



ADT - List - Positions

- A position p will be considered valid if it denotes the position of an actual element from the list:
 - if p is a pointer to a memory location, p is valid if it is the address of an element from a list (not NIL or some other address that is not the address of any element)
 - if *p* is the rank of the element from the list, *p* is valid if it is between 1 and the number of elements.
- ullet For an invalid position we will use the following notation: $oldsymbol{\perp}$



ADT List I

Domain of the ADT List:

 $\mathcal{L} = \{I | I \text{ is a list with elements of type TElem, each having a unique position in I of type TPosition} \}$

ADT List II

- init(I)
 - descr: creates a new, empty list
 - pre: true
 - **post:** $l \in \mathcal{L}$, l is an empty list

ADT List III

- first(I)
 - **descr:** returns the TPosition of the first element
 - pre: $l \in \mathcal{L}$
 - **post:** $first \leftarrow p \in TPosition$

$$p = egin{cases} ext{the position of the first element from I} & ext{if I}
eq \emptyset \ & ext{} & ext$$

ADT List IV

- last(l)
 - descr: returns the TPosition of the last element
 - pre: $I \in \mathcal{L}$
 - $\textbf{p ost: } \textit{last} \leftarrow p \in \textit{TPosition} \\ p = \begin{cases} \text{the position of the last element from I} & \text{if I} \neq \emptyset \\ \bot & \textit{otherwise} \end{cases}$

ADT List V

- valid(l, p)
 - descr: checks whether a TPosition is valid in a list
 - pre: $l \in \mathcal{L}, p \in TPosition$
 - **post:** $valid \leftarrow \begin{cases} true & \text{if p is a valid position in I} \\ false & otherwise \end{cases}$

ADT List VI

- next(I, p)
 - descr: goes to the next TPosition from a list
 - pre: $l \in \mathcal{L}, p \in TPosition$
 - post:

$$next \leftarrow q \in TPosition$$

```
q=
\left\{	ext{the position of the next element after p} \quad 	ext{if p is not the last position} 
ight.
\left.egin{aligned} oldsymbol{\perp} & \textit{otherwise} \end{aligned}
ight.
```

• throws: exception if p is not valid



ADT List VII

- previous(I, p)
 - descr: goes to the previous TPosition from a list
 - pre: $l \in \mathcal{L}, p \in TPosition$
 - post:

$$previous \leftarrow q \in TPosition$$

$$q = \begin{cases} \text{the position of the element before p} & \text{if p is not the first position} \\ \bot & \textit{otherwise} \end{cases}$$

throws: exception if p is not valid



ADT List VIII

- getElement(I, p, e)
 - descr: returns the element from a given TPosition
 - **pre**: $l \in \mathcal{L}, p \in TPosition$, valid(p)
 - **post:** $e \in TElem$, e = the element from position p from l
 - throws: exception if p is not valid

ADT List IX

- position(I, e)
 - descr: returns the TPosition of an element
 - pre: $l \in \mathcal{L}, e \in TElem$
 - post:

$$position \leftarrow p \in TPosition$$

$$p = \begin{cases} \text{the first position of element e from I} & \text{if } e \in I \\ \bot & \textit{otherwise} \end{cases}$$

ADT List X

- modify(I, p, e)
 - descr: replaces an element from a TPosition with another
 - **pre:** $l \in \mathcal{L}, p \in TPosition, e \in TElem$, valid(p)
 - **post:** $l' \in \mathcal{L}$, the element from position p from l' is e
 - throws: exception if p is not valid

ADT List XI

- insertFirst(I, e)
 - descr: inserts a new element at the beginning of a list
 - pre: $l \in \mathcal{L}, e \in TElem$
 - post: $l' \in \mathcal{L}$, the element e was added at the beginning of l

ADT List XII

- insertLast(I, e)
 - descr: inserts a new element at the end of a list
 - pre: $l \in \mathcal{L}, e \in TElem$
 - post: $l' \in \mathcal{L}$, the element e was added at the end of l

ADT List XIII

- insertAfter(I, p, e)
 - descr: inserts a new element after a given position
 - **pre:** $l \in \mathcal{L}, p \in TPosition, e \in TElem$, valid(p)
 - **post:** $l' \in \mathcal{L}$, the element e was added in l after the position p (position(l', e) = next(l', p) if e is not already in the list)
 - throws: exception if p is not valid

ADT List XIV

- insertBefore(I, p, e)
 - descr: inserts a new element before a given position
 - **pre:** $l \in \mathcal{L}, p \in TPosition, e \in TElem$, valid(p)
 - post: I' ∈ L, the element e was added in I before the position p (position(I', e) = previous(I', p) if e is not already in the list)
 - throws: exception if p is not valid

ADT List XV

- remove(I, p, e)
 - descr: removes an element from a given position from a list
 - **pre:** $l \in \mathcal{L}, p \in TPosition$, valid(p)
 - **post:** $e \in TElem$, e is the element from position p from I, $l' \in \mathcal{L}$, l' = I e.
 - throws: exception if p is not valid

ADT List XVI

- search(I, e)
 - descr: searches for an element in the list
 - pre: $l \in \mathcal{L}, e \in TElem$
 - post:

$$search \leftarrow \begin{cases} true & \text{if } e \in I \\ false & otherwise \end{cases}$$

ADT List XVII

- isEmpty(I)
 - descr: checks if a list is empty
 - pre: $l \in \mathcal{L}$
 - post:

$$isEmpty \leftarrow \begin{cases} true & \text{if } I = \emptyset \\ false & otherwise \end{cases}$$

ADT List XVIII

- size(I)
 - descr: returns the number of elements from a list
 - pre: $l \in \mathcal{L}$
 - **post:** *size* ← the number of elements from l

ADT List XIX

- destroy(I)
 - descr: destroys a list
 - pre: $l \in \mathcal{L}$
 - post: I was destroyed

ADT List XX

- iterator(I, it)
 - descr: returns an iterator for a list
 - pre: $l \in \mathcal{L}$
 - **post**: $it \in \mathcal{I}$, it is an iterator over I

TPosition

- Using TPositions in the interface of the ADT List can have disadvantages:
 - The exact type of a TPosition might differ if we use different representations for the list.
 - We have a large interface with many operations.

TPosition - C++

- In STL, TPosition is represented by an iterator.
- The operations valid, next, previous, getElement are actually operations for the iterator.
- For example vector:
 iterator insert(iterator position, const value_type& val)
 iterator erase (iterator position);
- For example list:
 iterator insert(iterator position, const value_type& val)
 iterator erase (iterator position);

TPosition - Java

- In Java, TPosition is represented by an index.
- We can add and remove using index and we can access elements using their index (but we have iterator as well for the List).
- There are fewer operations in the interface of the List
- For example:

```
void add(int index, E element)
E get(int index)
E remove(int index)
```



ADT SortedList

- We can define the ADT SortedList, in which the elements are memorized in a given order, based on a relation.
- Elements still have positions, we can access elements by position.
- Differences in the interface:
 - init takes as parameter a relation
 - only one insert operation exists
 - no modify operation

ADT List - representation

- If we want to implement the ADT List (or ADT SortedList)
 we can use the following data structures are representation:
 - a (dynamic) array elements are kept in a contiguous memory location we have direct access to any element
 - a linked list elements are kept in nodes, we do not have direct access to any element
- Demo



ADT Stack



Stack of books

Source: www.clipartfest.com

• The word stack might be familiar from expressions like: *stack* of books, *stack* of paper or from the *call stack* that you usually see in debug windows.

Stack II

- The ADT Stack represents a container in which access to the elements is restricted to one end of the container, called the top of the stack.
 - When a new element is added, it will automatically be added at the top.
 - When an element is removed it will be removed automatically from the top.
 - Only the element from the top can be accessed.
- Because of this restricted access, the stack is said to have a LIFO policy: Last In, First Out (the last element that was added will be the first element that will be removed).



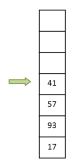
Stack III

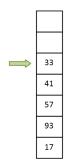
- When a new stack is created, it can have a fixed capacity. If the number of elements in the stack is equal to this capacity, we say that the stack is full.
- A stack with no elements is called an empty stack.

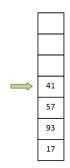
Stack Example

 Suppose that we have the following stack (green arrow shows the top of the stack): • We *push* the number 33:

• We pop an element:

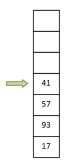






Stack Example II

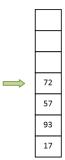
• This is our stack:



• We pop another element:



• We *push* the number 72:



Stack Interface I

- The domain of the ADT Stack: $S = \{s | s \text{ is a stack with elements of type TElem}\}$
- The interface of the ADT Stack contains the following operations:

Stack Interface II

- init(s)
 - Description: creates a new empty stack
 - Pre: True
 - **Post:** $s \in \mathcal{S}$, s is an empty stack

Stack Interface III

- destroy(s)
 - Description: destroys a stack
 - Pre: $s \in \mathcal{S}$
 - Post: s was destroyed

Stack Interface IV

- push(s, e)
 - **Description:** pushes (adds) a new element onto the stack
 - **Pre:** $s \in \mathcal{S}$, e is a *TElem*
 - **Post:** $s' \in \mathcal{S}$, $s' = s \oplus e$, e is the most recent element added to the stack
 - Throws: an overflow error if the stack is full

Stack Interface V

- pop(s)
 - Description: pops (removes) the most recent element from the stack
 - Pre: $s \in \mathcal{S}$
 - **Post:** $pop \leftarrow e$, e is a *TElem*, e is the most recent element from s, $s' \in S$, $s' = s \ominus e$
 - Throws: an underflow error if the stack is empty

Stack Interface VI

- top(s)
 - Description: returns the most recent element from the stack (but it does not change the stack)
 - Pre: $s \in \mathcal{S}$
 - Post: top ← e, e is a TElem, e is the most recent element from s
 - Throws: an underflow error if the stack is empty

Stack Interface VII

- isEmpty(s)
 - Description: checks if the stack is empty (has no elements)
 - Pre: $s \in \mathcal{S}$
 - Post:

$$isEmpty \leftarrow \left\{ egin{array}{ll} true, & if s has no elements \\ false, & otherwise \end{array} \right.$$

Stack Interface VIII

- isFull(s)
 - **Description:** checks if the stack is full not every representation has this operation
 - Pre: $s \in S$
 - Post:

$$isFull \leftarrow \left\{ egin{array}{l} true, \ if \ s \ is \ full \\ false, \ otherwise \end{array} \right.$$

Stack Interface IX

• **Note:** stacks cannot be iterated, so they don't have an *iterator* operation!

Representation for Stack

- Data structures that can be used to implement a stack:
 - Arrays
 - Static Array
 - Dynamic Array
 - Linked Lists
 - Singly-Linked List
 - Doubly-Linked List

Static Array-based representation II

? Where should we place the top of the stack for optimal performance?

Static Array-based representation II

? Where should we place the top of the stack for optimal performance?

- We have two options:
 - Place top at the beginning of the array every push and pop operation needs to shift every element to the right or left.
 - Place top at the end of the array push and pop elements without moving the other ones.

Static Array-based representation

Stack:

capacity: Integer

top: Integer

elements: TElem[]

Init - Implementation using a static array

```
subalgorithm init(s) is:
s.capacity \leftarrow MAX_CAPACITY
//MAX_CAPACITY is a constant with the maximum capacity
s.top \leftarrow 0
@allocate memory for the elements array
end-subalgorithm
```

• Complexity: $\Theta(1)$



Push - Implementation using a static array

```
\begin{array}{l} \textbf{subalgorithm} \ \ \text{push}(s, \, e) \ \textbf{is:} \\ \textbf{if} \ \ \text{s.capacity} = s.\text{top} \ \textbf{then} \ //\text{check} \ \textit{if} \ s \ \textit{is} \ \textit{full} \\ \text{ @throw overflow (full stack) exception} \\ \textbf{end-if} \\ \text{s.elements}[s.\text{top}+1] \leftarrow e \\ \text{s.top} \leftarrow s.\text{top} + 1 \\ \textbf{end-subalgorithm} \end{array}
```

• Complexity: $\Theta(1)$

Pop - Implementation using a static array

Complexity: Θ(1)

Top - Implementation using a static array

```
function top(s) is:
    if s.top = 0 then //check if s is empty
        @throw underflow(empty stack) exception
    end-if
    topElem ← s.elements[s.top]
    top← topElem
end-function
```

• Complexity: $\Theta(1)$

IsEmpty - Implementation using a static array

```
function isEmpty(s) is:

if s.top = 0 then

isEmpty ← True

else

isEmpty ← False

end-if

end-function
```

• Complexity: $\Theta(1)$

IsFull - Implementation using a static array

```
function isFull(s) is:

if s.top = s.capacity then

isFull ← True

else

isFull ← False

end-if

end-function
```

Implementation using a dynamic array

? Which operations change if we use a dynamic array instead of a static one for implementing a Stack?

Implementation using a dynamic array

? Which operations change if we use a dynamic array instead of a static one for implementing a Stack?

- If we use a Dynamic Array we can change the capacity of the Stack as elements are pushed onto it, so the Stack will never be full (except when there is no memory at all).
- The *push* operation does not throw an exception, it resizes the array if needed (doubles the capacity).
- The isFull operation will always return false.

Singly-Linked List-based representation

?Where should we place the top of the stack for optimal performance?

Singly-Linked List-based representation

?Where should we place the top of the stack for optimal performance?

- We have two options:
 - Place it at the end of the list (like we did when we used an array) - for every push, pop and top operation we have to iterate through every element to get to the end of the list.
 - Place it at the beginning of the list we can push and pop elements without iterating through the list.

Singly-Linked List-based representation

Node:

elem: TElem next: ↑ Node

Stack

top: ↑ Node

Init - Implementation using a singly-linked list

$\begin{array}{l} \textbf{subalgorithm} \ \, \textbf{init}(\textbf{s}) \ \, \textbf{is:} \\ \text{s.top} \ \, \leftarrow \ \, \textbf{NIL} \\ \textbf{end-subalgorithm} \end{array}$

Destroy - Implementation using a singly-linked list

```
subalgorithm destroy(s) is:

while s.top ≠ NIL execute

firstNode ← s.top

s.top ← [s.top].next

@deallocate firstNode

end-while

end-subalgorithm
```

• Complexity: $\Theta(n)$ - where n is the number of elements from s

Push - Implementation using a singly-linked list

```
subalgorithm push(s, e) is:
   //allocate a new Node and set its fields
   @allocate newnode of type Node
   [newnode].elem \leftarrow e
   [newnode].next \leftarrow NIL
   if s.top = NIL then
      s.top \leftarrow newnode
   else
      [newnode].next \leftarrow s.top
      s.top \leftarrow newnode
   end-if
end-subalgorithm
```

Pop - Implementation using a singly-linked list

```
function pop(s) is:
  if s.top = NIL then //check if s is empty
     Othrow underflow(empty stack) exception
  end-if
  firstNode \leftarrow s.top
  topElem \leftarrow [firstNode].elem
  s.top \leftarrow [s.top].next
  @deallocate firstNode
  pop← topElem
end-function
```

Top - Implementation using a singly-linked list

```
function top(s) is:
    if s.top = NIL then //check if s is empty
        @throw underflow(empty stack) exception
    end-if
    topElem ← [s.top].elem
    top← topElem
end-function
```

IsEmpty - Implementation using a singly-linked list

```
function isEmpty(s) is:

if s.top = NIL then

isEmpty ← True

else

isEmpty ← False

end-if

end-function
```

IsFull - Implementation using a singly-linked list

 We don't have a maximum capacity in case of a linked list, so our stack will never be full. If we still want to implement this method, we can make it to always return false.

```
function isFull(s) is:
isFull ← False
end-function
```



Fixed capacity stack with singly-linked list

? How could we implement a stack with a fixed maximum capacity using a singly-linked list?

Fixed capacity stack with singly-linked list

? How could we implement a stack with a fixed maximum capacity using a singly-linked list?

 Similar to the implementation with a static array, we can keep in the Stack structure two integer values: maximum capacity and current size.

GetMinimum in constant time

? How can we design a *special stack* that has a *getMinimum* operation with $\Theta(1)$ time complexity (and the other operations have $\Theta(1)$ time complexity as well)?

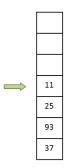
GetMinimum in constant time

? How can we design a *special stack* that has a *getMinimum* operation with $\Theta(1)$ time complexity (and the other operations have $\Theta(1)$ time complexity as well)?

 We can keep an auxiliary stack, containing as many elements as the original stack, but containing the minimum value up to each element. Let's call this auxiliary stack a min stack and the original stack the element stack.

GetMinimum in constant time - Example

• If this is the *element stack*:



• This is the corresponding *min stack*:

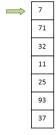


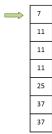
GetMinimum in constant time - Example

When a new element is pushed to the element stack, we push
a new element to the min stack as well. This element is the
minimum between the top of the min stack and the newly
added element.

The element stack:

• The corresponding *min stack*:





GetMinimum in constant time

 When an element si popped from the element stack, we will pop an element from the min stack as well.

 The getMinimum operation will simply return the top of the min stack.

 The other stack operations remain unchanged (except init, where you have to create two stacks).

GetMinimum in constant time

• Let's implement the *push*, *pop* and *getMinimum* operations for this *SpecialStack*, represented in the following way:

SpecialStack:

elementStack: Stack minStack: Stack

 We will use an existing implementation for the stack and work only with the operations from the interface.

Push for SpecialStack

```
subalgorithm push(ss, e) is:
  if isFull(ss.elementStack) then
     Othrow overflow (full stack) exception
  end-if
  if isEmpty(ss.elementStack) then//the stacks are empty, just push the elem
     push(ss.elementStack, e)
     push(ss.minStack, e)
  else
     push(ss.elementStack, e)
     currentMin ← top(ss.minStack)
     if currentMin < e then //find the minim to push to minStack
        push(ss.minStack, currentMin)
     else
        push(ss.minStack, e)
     end-if
  end-if
end-subalgorithm //Complexity: \Theta(1)
```

Pop for SpecialStack

```
function pop(ss) is:
    if isEmpty(ss.elementStack) then
        @throw underflow (empty stack) exception
    end-if
    currentElem ← pop(ss.elementStack)
    pop(ss.minStack) //we don't need the value, just to pop it
    pop ← currentElem
end-function
```

GetMinimum for SpecialStack

```
function getMinimum(ss) is:
   if isEmpty(ss.elementStack) then
      @throw underflow (empty stack) exception
   end-if
   getMinimum ← top(ss.minStack)
end-function
```

Complexity: Θ(1)

SpecialStack - Notes / Think about it

- We designed the special stack in such a way that all the operations have a $\Theta(1)$ time complexity.
- The disadvantage is that we occupy twice as much space as with the regular stack.

? Think about how can we reduce the space occupied by the *min stack* to O(n) (especially if the minimum element of the stack rarely changes). *Hint: If the minimum does not change, we don't have to push a new element to the min stack.* How can we implement the *push* and *pop* operations in this case? What happens if the minimum element appears more than once in the *element stack*?