6.2. An abstract datatype for stacks

Stacks of objects of type $\alpha$: $\alpha$ stack

Operations

value emptyStack
TYPE: $\alpha$ stack
VALUE: the empty stack

function isEmptyStack S
TYPE: $\alpha$ stack $\rightarrow$ bool
PRE: (none)
POST: true if S is empty
    false otherwise

function push v S
TYPE: $\alpha$ $\rightarrow$ $\alpha$ stack $\rightarrow$ $\alpha$ stack
PRE: (none)
POST: the stack S with v added as new top element

function top S
TYPE: $\alpha$ stack $\rightarrow$ $\alpha$
PRE: S is non-empty
POST: the top element of S

function pop S
TYPE: $\alpha$ stack $\rightarrow$ $\alpha$ stack
PRE: S is non-empty
POST: the stack S without its top element
6.3. Realisation of the stack abstract datatype

Version 1

Representation of a stack by a list:

```
type α stack = α list
```

REPRESENTATION CONVENTION: the head of the list is the top of the stack, the 2nd element of the list is the element below the top, etc.

Realisation of the operations (stack1.sml)

```
val emptyStack = [ ]
fun isEmptyStack S = (S = [ ])
fun push v S = v::S
fun top [ ] = error "top: empty stack"
   | top (x::xs) = x
fun pop [ ] = error "pop: empty stack"
   | pop (x::xs) = xs
```

- This realisation does not force the usage of the stack type
- The operations can also be used with objects of type α list, even if they do not represent stacks!
- It is possible to access the elements of the stack without using the operations specified above: no encapsulation!
Version 2

Definition of a new constructed type using the list type:

```plaintext
datatype α stack = Stack of α list
```

REPRESENTATION CONVENTION: the head of the list is the top of the stack, the 2nd element of the list is the element below the top, etc

Realisation of the operations

```plaintext
val emptyStack = Stack [ ]

fun isEmptyStack (Stack S) = (S = [ ])

fun push v (Stack S) = Stack (v::S)

fun top (Stack [ ]) = error "top: empty stack"
   | top (Stack (x::xs)) = x

fun pop (Stack [ ]) = error "pop: empty stack"
   | pop (Stack (x::xs)) = Stack xs
```

- The operations are now only defined for stacks
- It is still possible to access the elements of the stack without using the operations specified above, namely by pattern matching
An abstract datatype (stack2.sml)

Objective: encapsulate the definition of the stack type and its operations in a parameterised abstract datatype

abstype 'a stack = Stack of 'a list

val emptyStack = Stack [ ]

fun isEmptyStack (Stack S) = (S = [ ])

fun push v (Stack S) = Stack (v::S)

fun top (Stack [ ]) = error "top: empty stack"
   | top (Stack (x::xs)) = x

fun pop (Stack [ ]) = error "pop: empty stack"
   | pop (Stack (x::xs)) = Stack xs

end

• The stack type is an abstract datatype (ADT)
• The concrete representation of a stack is hidden
• An object of the stack type can only be manipulated via the functions defined in its ADT declaration
• The Stack constructor is invisible outside the ADT
• It is now impossible to access the representation of a stack outside the declarations of the functions of the ADT
• The parameterisation allows the usage of stacks of integers, reals, strings, integer functions, etc, from a single definition!
- `abstype 'a stack = Stack of 'a list with ... ;
  type 'a stack
  val 'a emptyStack = - : 'a stack
  val ''a isEmptyStack = fn : ''a stack -> bool
  ...
- push 1 (Stack [ ]) ;
  Error: unbound variable or constructor: Stack
- push 1 emptyStack ;
  val it = - : int stack

It is *impossible* to compare two stacks:
- emptyStack = emptyStack ;
  Error: operator and operand don’t agree
  [equality type required]

It is *impossible* to see the contents of a stack without popping its elements, so let us add a visualisation function:

```plaintext
function showStack S
  TYPE: α stack → α list
  PRE: (none)
  POST: the representation of S in list form, with the top of S as head, etc

abstype 'a stack = Stack of 'a list
with

  ...
  fun showStack (Stack S) = S
end
```

- The result of `showStack` is *not* of the `stack` type
- One can thus *not* apply the stack operations to it
Version 3

Definition of a recursive new constructed type:

```
datatype α stack = EmptyStack

           | >> of α stack * α
```

**EXAMPLE:** EmptyStack >> 3 >> 5 >> 2 represents the stack with top 2

**REPRESENTATION CONVENTION:** the right-most value is the top of
the stack, its left neighbour is the element below the top, etc

*An abstract datatype* (stack3.sml)

```
abstype 'a stack = EmptyStack | >> of 'a stack * 'a

with

  infix >>

  val   emptyStack = EmptyStack

  fun   isEmptyStack EmptyStack   = true

  fun   push v S = S>>v

  fun   top EmptyStack = error "top: empty stack"

  fun   pop EmptyStack = error "pop: empty stack"

  fun   showStack EmptyStack = [ ]

end
```

We have thus defined a new list constructor,
but with access to the elements *from the right*!
6.4. An abstract datatype for FIFO queues

First-in first-out (FIFO) queues of objects of type $\alpha$: $\alpha$ queue

- Addition of elements to the rear ($tail$)
- Deletion of elements from the front ($head$)

Operations

**value** emptyQueue
TYPE: $\alpha$ queue
VALUE: the empty queue

**function** isEmptyQueue $Q$
TYPE: $\alpha$ queue $\rightarrow$ bool
PRE: (none)
POST: true if $Q$ is empty
false otherwise

**function** enqueue $v$ $Q$
TYPE: $\alpha$ $\rightarrow$ $\alpha$ queue $\rightarrow$ $\alpha$ queue
PRE: (none)
POST: the queue $Q$ with $v$ added as new tail element

**function** head $Q$
TYPE: $\alpha$ queue $\rightarrow$ $\alpha$
PRE: $Q$ is non-empty
POST: the head element of $Q$
function dequeue Q
TYPE: $\alpha$ queue $\rightarrow$ $\alpha$ queue
PRE: Q is non-empty
POST: the queue Q without its head element

function showQueue Q
TYPE: $\alpha$ queue $\rightarrow$ $\alpha$ list
PRE: (none)
POST: the representation of Q in list form, with the head of Q as head, etc

‘Formal’ semantics

isEmptyQueue emptyQueue = true
$\forall v, Q :$ isEmptyQueue (enqueue v Q) = false
head emptyQueue = ... error ...
$\forall v, Q :$ head (enqueue v Q) = if isEmptyQueue Q then v else head Q
dequeue emptyQueue = ... error ...
$\forall v, Q :$ dequeue (enqueue v Q) = if isEmptyQueue Q then emptyQueue else enqueue v (dequeue Q)
6.5. Realisation of the queue abstract datatype

Version 1

Representation of a FIFO queue by a list:

\[\text{type } \alpha \text{ queue } \equiv \alpha \text{ list}\]

REPRESENTATION CONVENTION: the head of the list is the head of the queue, the 2nd element of the list is behind the head of the queue, and so on, and the last element of the list is the tail of the queue.

Example: the queue

\[
\begin{array}{cc}
\text{head} & \text{tail} \\
3 & 87502
\end{array}
\]

is represented by the list \([3,8,7,5,0,2]\)

Exercises

- Realise the queue ADT using this representation
- What is the time complexity of enqueuing an element?
- What is the time complexity of dequeuing an element?
Version 2

Representation of a FIFO queue by a pair of lists:

```
datatype α queue = Queue of α list * α list
```

REPRESENTATION CONVENTION: the term

```
Queue ([x₁, x₂, ..., xₙ], [y₁, y₂, ..., yₘ])
```

represents the queue

```
x₁ x₂ ... xₙ yₘ ... y₂ y₁
```

REPRESENTATION INVARIANT: (see next slide)

- It is now possible to enqueue in Θ(1) time
- It is still possible to dequeue in Θ(1) time, but only if \( n \geq 1 \)
- What if \( n = 0 \) while \( m > 0 \)?!
- The same queue can thus be represented in different ways
- How to test the equality of two queues?
Normalisation

Objective: avoid the case where \( n = 0 \) while \( m > 0 \)

When this case appears, transform (or: normalise) the representation of the queue:
transform \( \text{Queue} ([\ ], [y_1, \ldots, y_m]) \) with \( m > 0 \)
into \( \text{Queue} ([y_m, \ldots, y_1], [\ ]) \),
which indeed represents the same queue

We thus have:

\text{REPRESENTATION INVARIANT: a non-empty queue is never represented by} \ \text{Queue} ([\ ], [y_1, \ldots, y_m])

\textbf{function} normalise \( Q \)
\textbf{TYPE:} \( \alpha \) queue \( \rightarrow \alpha \) queue
\textbf{PRE:} (none)
\textbf{POST:} if \( Q \) is of the form \( \text{Queue} ([\ ], [y_1, \ldots, y_m]) \)
\hspace{1em} then \( \text{Queue} ([y_m, \ldots, y_1], [\ ]) \)
\hspace{1em} else \( Q \)

\textit{Realisation of the operations} (queue2.sm1)

Construction of an abstract datatype:
the \texttt{normalise} function may be \textit{local} to the ADT,
as it is only used for realising some operations on queues
Ch.6: Abstract Datatypes

6.5. Realisation of the queue abstract datatype

abstype 'a queue = Queue of 'a list * 'a list

with

val emptyQueue = Queue ([ ],[ ])

fun isEmptyQueue (Queue ([ ],[ ])) = true
 | isEmptyQueue (Queue (xs,ys)) = false

fun head (Queue (x::xs,ys)) = x
 | head (Queue ([ ],[ ])) = error "head: empty queue"
 | head (Queue ([ ],y::ys)) = error "head: non-normalised queue"

local

fun normalise (Queue ([ ],ys)) = Queue (rev ys,[ ])
 | normalise Q = Q

in

fun enqueue v (Queue (xs,ys)) = normalise (Queue (xs,v::ys))

fun dequeue (Queue (x::xs,ys)) = normalise (Queue (xs,ys))
 | dequeue (Queue ([ ],[ ])) = error "dequeue: empty queue"
 | dequeue (Queue ([ ],y::ys)) = error "dequeue: non-norm. queue"

end

fun showQueue (Queue (xs,ys)) = xs @ (rev ys)

fun equalQueues Q1 Q2 = (showQueue Q1 = showQueue Q2)

end

• Why do the head and dequeue functions not normalise the queue instead of stopping the execution with an error?

• The normalisation and representation invariant are hidden in the realisation of the abstract datatype

• On average, the time of enqueueing and dequeuing is $\Theta(1)$

• This representation is thus very efficient!
6.6. An abstract datatype for binary trees

Concepts and terminology

Binary trees of objects of type $\alpha$: $\alpha \ bTree$

Operations

**value** emptyBtree
TYPE: $\alpha \ bTree$
VALUE: the empty binary tree

**function** isEmptyBtree T
TYPE: $\alpha \ bTree \rightarrow bool$
PRE: (none)
POST: 
true if T is empty
false otherwise
function consBtree v L R
TYPE: $\alpha \rightarrow \alpha$ bTree $\rightarrow \alpha$ bTree $\rightarrow \alpha$ bTree
PRE: (none)
POST: the binary tree with root v, left sub-tree L, and right sub-tree R

function left T
TYPE: $\alpha$ bTree $\rightarrow \alpha$ bTree
PRE: T is non-empty
POST: the left sub-tree of T

function right T
TYPE: $\alpha$ bTree $\rightarrow \alpha$ bTree
PRE: T is non-empty
POST: the right sub-tree of T

function root T
TYPE: $\alpha$ bTree $\rightarrow \alpha$
PRE: T is non-empty
POST: the root of T

'Formal' semantics

isEmptyBtree emptyBtree = true
$\forall v, L, R :$ isEmptyBtree (consBtree v L R) = false

root emptyBtree = ... error ...
$\forall v, L, R :$ root (consBtree v L R) = v

left emptyBtree = ... error ...
$\forall v, L, R :$ left (consBtree v L R) = L

right emptyBtree = ... error ...
$\forall v, L, R :$ right (consBtree v L R) = R
6.7. Realisation of the bTree abstract datatype

Representation

datatype bTree = Void
  | Bt of int * bTree * bTree

REPRESENTATION CONVENTION: a binary tree with root x, left subtree L, and right subtree R is represented by Bt(x,L,R)
EXAMPLE: Bt(4, Bt(2, Bt(1,Void,Void), Bt(3,Void,Void)), Bt(8, Bt(6, Bt(5,Void,Void), Bt(7,Void,Void)), Bt(9,Void,Void)))

Realisation of the operations (bTree.sml)

abstype 'a bTree = Void
  | Bt of 'a * 'a bTree * 'a bTree

with

  val emptyBtree = Void
  fun isEmptyBtree Void = true
    | isEmptyBtre (Bt(v,L,R)) = false
  fun consBtree v L R = Bt(v,L,R)
  fun left Void = error "left: empty bTree"
    | left (Bt(v,L,R)) = L
  fun right Void = error "right: empty bTree"
    | right (Bt(v,L,R)) = R
  fun root Void = error "root: empty bTree"
    | root (Bt(v,L,R)) = v

end
Walk operations (inorder.sml)

function inorder T
TYPE: $\alpha$ bTree $\rightarrow$ $\alpha$ list
PRE: (none)
POST: the nodes of T upon an inorder walk

fun inorder Void = [ ]
  | inorder (Bt(v,L,R)) = (inorder L) @ (v :: inorder R)

No tail recursion!
It takes $\Theta(n \log n)$ time for a binary tree of $n$ nodes...

function inorderGen T acc
TYPE: $\alpha$ bTree $\rightarrow$ $\alpha$ list $\rightarrow$ $\alpha$ list
PRE: (none)
POST: (the nodes of T upon an inorder walk) @ acc

fun inorderGen Void acc = acc
  | inorderGen (Bt(v,L,R)) acc =
    let val rAcc = inorderGen R acc
    in inorderGen L (v::rAcc) end

fun inorder t = inorderGen t [ ]

One tail recursion! No call to @ (concatenation)!
It takes $\Theta(n)$ time for a binary tree of $n$ nodes

Exercises

• Efficiently realise the preorder and postorder walks
  of a binary tree, and analyse the underlying algorithms

• How to test the equality of two binary trees?
Other operations

function exists k T
TYPE: $\alpha \rightarrow \alpha \rightarrow$ bTree $\rightarrow$ bool
PRE: (none)
POST: true if T contains node k
    false otherwise

function insert k T
TYPE: $\alpha \rightarrow \alpha \rightarrow$ bTree $\rightarrow$ $\alpha \rightarrow$ bTree
PRE: (none)
POST: T with node k

function delete k T
TYPE: $\alpha \rightarrow \alpha \rightarrow$ bTree $\rightarrow$ $\alpha \rightarrow$ bTree
PRE: (none)
POST: if exists k T, then T without one occurrence of node k, otherwise T

function nbNodes T
TYPE: $\alpha$ bTree $\rightarrow$ int
PRE: (none)
POST: the number of nodes of T

function nbLeaves T
TYPE: $\alpha$ bTree $\rightarrow$ int
PRE: (none)
POST: the number of leaves of T

Exercises

• Efficiently realise these five functions

• Show that their algorithms at worst take $\Theta(n)$ time,
  if not $\Theta(1)$ time, on a binary tree with initially $n$ nodes
Height of a binary tree (height.sml)

- The *height of a node* is the length of the longest path (measured in its number of nodes) from that node to a leaf.

- The *height of a tree* is the height of its root.

function height T
TYPE: \( \alpha \) bTree \( \rightarrow \) int
PRE: (none) ; POST: the height of T

fun height Void = 0
| height (Bt(v,L,R)) = 1 + Int.max (height L, height R)

No tail recursion!
It takes \( \Theta(n) \) time for a binary tree of \( n \) nodes.
Note that heightGen T acc = acc + height T does not suffice to get a tail recursion: why?!

function heightGen T acc hMax
TYPE: \( \alpha \) bTree \( \rightarrow \) int \( \rightarrow \) int \( \rightarrow \) int
PRE: (none) ; POST: max (acc + height T, hMax)

fun heightGen Void acc hMax = Int.max (acc, hMax)
| heightGen (Bt(v,L,R)) acc hMax =
  let val h1 = heightGen R (acc+1) hMax
  in heightGen L (acc+1) h1 end

fun height2 bt = heightGen bt 0 0

One tail recursion!
It also takes \( \Theta(n) \) time for a binary tree of \( n \) nodes, but it takes less space!
6.8. An ADT for Binary Search Trees (BSTs)

Concepts and terminology (see the tree on slide 6.19)

Binary search trees of nodes of type $(\alpha, \beta)$: $(\alpha, \beta) \text{bsTree}$
where:

- $\alpha$ is the type of the keys (need for an equality test)
- $\beta$ is the type of the satellite data for each key

are binary trees with:

(REPRESENTATION) INVARIANT: for a binary search tree with (k,s) in the root, left subtree L, and right subtree R:
- every element of L has a key smaller than k
- every element of R has a key larger than k

Note that we (arbitrarily) ruled out duplicate keys

Benefits

- The inorder walk of a binary search tree lists its nodes by increasing order of their keys!
- The basic operations at worst take $\Theta(n)$ time on a binary search tree with (initially) $n$ nodes, but they take $O(\lg n)$ time on randomly built binary search trees

Let us restrict our realisation to integer keys: $\beta \text{bsTree}$
Some operations

value emptyBsTree
TYPE: β bsTree
VALUE: the empty binary search tree

function isEmptyBsTree T
TYPE: β bsTree → bool
PRE: (none)
POST: true if T is empty
     false otherwise

function exists k T
TYPE: int → β bsTree → bool
PRE: (none)
POST: true if T contains a node with key k
     false otherwise

function insert (k,s) T
TYPE: (int × β) → β bsTree → β bsTree
PRE: (none)
POST: if exists k T, then T with s as satellite data for key k
     otherwise T with node (k,s)

function retrieve k T
TYPE: int → β bsTree → β
PRE: exists k T
POST: the satellite data associated to key k in T

function delete k T
TYPE: int → β bsTree → β bsTree
PRE: (none)
POST: if exists k T, then T without the node with key k, otherwise T
6.9. Realisation of the bsTree ADT

Representation

datatype 'b bsTree = Void
  | Bst of (int * 'b) * 'b bsTree * 'b bsTree

REPRESENTATION CONVENTION: a BST with (k,s) in the root, left subtree L, and right subtree R is represented by Bst((k,s),L,R)
REPRESENTATION IN Variant: (see slide 6.25)

Realisation of the operations (bsTree.sml)

val emptyBsTree = Void

fun isEmptyBsTree Void = true
  | isEmptyBsTree (Bst((key,sat),L,R)) = false

fun exists k Void = false
  | exists k (Bst((key,sat),L,R)) =
    if k = key then true
    else if k < key then exists k L
    else (k > key *) exists k R

fun insert (k,s) Void = Bst((k,s),Void,Void)
  | insert (k,s) (Bst((key,sat),L,R)) =
    if k = key then Bst((k,s),L,R)
    else if k < key then Bst((key,sat), (insert (k,s) L), R)
    else (k > key *) Bst((key,sat), L, (insert (k,s) R))

fun retrieve k Void = error "retrieve: non-existing node"
  | retrieve k (Bst((key,sat),L,R)) =
    if k = key then sat
    else if k < key then retrieve k L
    else (k > key *) retrieve k R
When deleting a node \((\text{key}, \text{sat})\) whose subtrees \(L\) and \(R\) are both non-empty, we must not violate the repr. invariant!

1. Replace \((\text{key}, \text{sat})\) by the node with the \textit{maximal} key of \(L\), whose key is smaller than the key of any node of \(R\) (one could also replace by the node with the \textit{minimal} key of \(R\))

2. Remove this node with the maximal key from \(L\)

So we need a \texttt{deleteMax} function:

\begin{verbatim}
function deleteMax T
  TYPE: \(\beta\) bsTree \(\rightarrow\) (int * \(\beta\)) * \(\beta\) bsTree
  PRE: T is non-empty
  POST: (max, NT), where max is the node of T with the maximal key, and NT is T without max

fun deleteMax Void = error "deleteMax: empty bsTree"
| deleteMax (Bst(r,L,Void): 'b bsTree) = (r, L)
| deleteMax (Bst(r,L,R)) =
  let val (max, newR) = deleteMax R
  in (max, Bst(r,L,newR)) end

fun delete k Void = Void
| delete k (Bst((key,sat),L,R)) =
  if k < key then Bst((key,sat), (delete k L), R)
  else if k > key then Bst((key,sat), L, (delete k R))
  else (* k = key *)
    case (L,R) of
      (Void, _ ) => R
      ( _ ,Void) => L
      ( _ , _ ) => let val (max, newL) = deleteMax L
                    in Bst(max,newL,R) end

end
\end{verbatim}