Exam Questions
Proof Methods for Concurrent Programs

11 February 2010

Instructions: only printed documents are authorised. You can admit the result of one question and move on. Leave optional questions until the end.

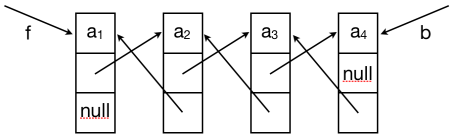
Your goal is to implement and prove correct some support functions used by the *A-Steal* multiprocessor scheduling algorithm.

Exercise 1. (Data structures)

A *double ended queue* (often abbreviated to *deque*) is a data structure that implements a queue for which elements can be added to or removed from the front (head) or back (tail). We follow the C++ interface (`std::deque`) to name the different operations:

`push_back`	insert element at back
`push_front`	insert element at front
`pop_back`	remove element from back
`pop_front`	remove element from front

We use *doubly-linked lists* to implement deques. A doubly-linked list can be graphically represented as



where `f` and `b` are the pointers to the front and back elements. When the doubly-linked list is empty, `f` = `b` = `null`. An implementation of the `push_front` procedure is given below:

```
push_front(a,f,b) {
  t := new (a,f,null);
  if (f = null) then b := t else [f+2] := t;
  f := t;
}
```

Remark that our notation for procedures assumes that the arguments (e.g. `a`, `f`, `b`) are passed by reference (in other terms, they can be considered global variables), while the other variables (e.g. `t`) are local.

1. Give a sequential implementation to the `pop_back(f,b,r)` and `pop_front(f,b,r)` procedures:

   - `f` and `b` are as above, while `r` is used to return either the popped value or `null` if the deque is empty;
   - your code should not leak memory.

Let α range over lists of values (ϵ denotes the empty list and $\cdot$ is the concatenation operator). Recall the recursive specification of doubly-linked list segments:

$$\text{dlseg } \epsilon\ (\texttt{f}, \texttt{f'}, \texttt{b'}, \texttt{b}) \ = \ \textit{empty} \ \wedge \ \texttt{f} = \texttt{b'} \ \wedge \ \texttt{f'} = \texttt{b}$$

$$\text{dlseg } (\texttt{a} \cdot \alpha)\ (\texttt{f}, \texttt{f'}, \texttt{b'}, \texttt{b}) \ = \ \exists \texttt{j}.\, \texttt{f} \mapsto \texttt{a}, \texttt{j}, \texttt{f'} \ * \ \text{dlseg } \alpha\ (\texttt{j}, \texttt{f}, \texttt{b'}, \texttt{b})$$

and consider the definition of doubly-linked lists below:

$$\text{dls } \alpha\ (\texttt{f}, \texttt{b}) \ = \ \text{dlseg } \alpha\ (\texttt{f}, \texttt{null}, \texttt{null}, \texttt{b})$$

Intuitively dls α (`f`, `b`) should be read "`f`, `b` are the ends of a doubly-linked list representing the list of values α".

2. Prove the following properties:

$$(a.) \qquad \text{dls } (\mathtt{a} \cdot \epsilon)\ (\mathtt{f}, \mathtt{b}) \ \Leftrightarrow\ \mathtt{f} \mapsto \mathtt{a}, \mathtt{null}, \mathtt{null} \wedge \mathtt{b} = \mathtt{f}$$

$$(b.) \qquad \text{dls } \alpha\ (\mathtt{f}, \mathtt{b}) \ \Rightarrow\ (\mathtt{f} = \mathtt{null} \Rightarrow (\alpha = \epsilon \wedge \mathtt{b} = \mathtt{null}))$$

$$(c.) \qquad \text{dls } \alpha\ (\mathtt{f}, \mathtt{b}) \ \Rightarrow\ (\mathtt{f} \neq \mathtt{null} \Rightarrow (\exists \mathtt{a}, \alpha'.\ \alpha = \mathtt{a} \cdot \alpha'))$$

3. Prove that the operations on deques satisfy the specifications below:

(*a.*)	$\{\text{dls } \alpha\ (\mathtt{f}, \mathtt{b})\}$	`push_front(a,f,b)`	$\{\text{dls } (\mathtt{a} \cdot \alpha)\ (\mathtt{f}, \mathtt{b})\}$
(*b.*)	$\{\text{dls } \epsilon\ (\mathtt{f}, \mathtt{b})\}$	`pop_front(f,b,r)`	$\{\text{dls } \epsilon\ (\mathtt{f}, \mathtt{b}) \wedge \mathtt{r} = \mathtt{null}\}$
(*c.*)	$\{\text{dls } (\mathtt{a} \cdot \alpha)\ (\mathtt{f}, \mathtt{b})\}$	`pop_front(f,b,r)`	$\{\text{dls } \alpha\ (\mathtt{f}, \mathtt{b}) \wedge \mathtt{r} = \mathtt{a}\}$
(*d. optional*)	$\{\text{dls } \epsilon\ (\mathtt{f}, \mathtt{b})\}$	`pop_back(f,b,r)`	$\{\text{dls } \epsilon\ (\mathtt{f}, \mathtt{b}) \wedge \mathtt{r} = \mathtt{null}\}$
(*e. optional*)	$\{\text{dls } (\alpha \cdot \mathtt{a})\ (\mathtt{f}, \mathtt{b})\}$	`pop_back(f,b,r)`	$\{\text{dls } \alpha\ (\mathtt{f}, \mathtt{b}) \wedge \mathtt{r} = \mathtt{a}\}$

4. Recall the definition of *precise predicate*. Is dlseg α (`f,f',b'`, `b`) a precise predicate?

In what follows you can assume that dls α (`f`, `b`) is precise.

Exercise 2. (The *A-Steal* scheduler)

The *A-Steal* algorithm implements task scheduling for several processors. A separate deque with the thread-identifiers (`tid`s) of the threads to be executed is maintained for each processor. To execute the next thread, the processor gets the first element from the deque (using the `pop_front` deque operation). If the current thread forks, it is put back to the front of the deque (`push_front`) and a new thread is executed. When one of the processors finishes execution of its own threads (i.e. its deque is empty), it can *steal* a thread from another processor: it gets the last element from the deque of another processor (`pop_back`) and executes it.

We introduce a handy C-like notation for arrays: if `a` points to a series of n contiguous memory locations, we write `a[i]` for `[a+i]`.

Consider a system with n processors, indexed from 0 to $n-1$. To implement the A-Steal algorithm we use two arrays, named `a` and `l`, and n deques. The array `a` has $2n$ entries, and `a[2*i]` and `a[2*i+1]` are the front and back addresses of the deque associated to the processor `i`. The elements of the deques are `tid`s (you can assume that a `tid` is just an integer). The array `l` has n entries, and each `l[i]` is a resource that protects the locations `a[2*i]` and `a[2*i+1]` and the associated deque.

1. Implement the `fork` and `schedule` procedures of the A-Steal algorithm, according to their informal specification below:

   `fork(proc,tid)` stores `tid` at the front of the deque of processor `proc`;

   `schedule(proc):tid` pops (and returns) the `tid` at the front of the deque of processor `proc`; if the deque of processor `proc` is empty, it pops (and returns) the `tid` at the back of the deque of another processor which has a non-empty deque.

   The procedures `fork` and `schedule` should not fail (`schedule` might loop), and, needless to say, should be robust to concurrent invocations.

2. Define the resource invariant associated to each resource `l[i]`.

3. Prove that

$$(a.) \qquad \{0 \leq \mathtt{proc} \leq n-1\}\ \mathtt{fork(proc,tid)}\ \{\mathit{true}\}$$

$$(b.) \qquad \{0 \leq \mathtt{proc} \leq n-1\}\ \mathtt{schedule(proc)}\ \{\mathtt{tid} \neq \mathtt{null}\}$$