Formal Methods in Software Development
Exercise 4 (July 11)

Wolfgang Schreiner
Wolfgang.Schreiner@risc.uni-linz.ac.at

June 3, 2005

The result is to me submitted to me by **July 11** (hard deadline) as an email that includes as attachments for each exercise the PVS file (.pvs) and the corresponding proof file (.prf).

Questions can be asked per email or in the classes before the deadline.

1 PVS Training

On the Web site you find a PVS file "exercise.pvs". Use PVS to prove the lemmas A, B, C, D, E, T, S in this file.

The lemmas A–E are simple predicate-logic proofs that only require the PVS commands `skolem!`, `inst`, `flatten`, `split`, and `assert`. The use of the `grind` command is *not* allowed.

Lemma T can be proved by an additional (single) use of the `expand` command (please note that `(expand` *name num pos* `)` expands occurence number *pos* of name *name* in formula number *num*).

Lemma S can be proved by induction on one of its arguments (command `induct`).

I strongly suggest that you try these proofs as soon as possible (not tomorrow, NOW!) such that you get familar with PVS and can solve the following exercise.

2 Binary Search

Take the Hoare triple

```
{$olda = a \land oldx = x \land (\forall i : 0 \le i < length(a) - 1 \Rightarrow a[i] \le a[i+1]) \land$
  $r = -1 \land low = 0 \land high = length(a) - 1$}
**while** $r = -1 \land low \le high$ **do**
  $mid := \lfloor (low + high)/2 \rfloor$
  **if** $a[mid] = x$ **then**
    $r := mid$ // B1
  **else if** $a[mid] < x$ **then**
    $low = mid + 1$ // B2
  **else**
    $high = mid - 1$ // B3
  **end**
**end**
{$a = olda \land x = oldx \land$
  $((r = -1 \land (\forall i : 0 \le i < length(a) \Rightarrow a[i] \ne x)) \lor$
   $(0 \le r < length(a) \land a[r] = x))$}
```

which describes the core proof obligation for a program that applies the "binary search" method to determine the index r of an element x in a sorted array a.

This Hoare triple can be verified with the help of the following loop invariant (take your time to understand especially the range conditions on *low* and *high*):

$$\begin{aligned}
&olda = a \land oldx = x \land (\forall i : 0 \le i < length(a) - 1 \Rightarrow a[i] \le a[i+1]) \land \\
&-1 \le r < length(a) \land (r \ne -1 \Rightarrow a[r] = x) \land \\
&0 \le low \le length(a) \land -1 \le high < length(a) \land low \le high + 1 \land \\
&(\forall i : 0 \le i < length(a) \land i < low \Rightarrow a[i] < x) \land \\
&(\forall i : 0 \le i < length(a) \land high < i \Rightarrow x < a[i])
\end{aligned}$$

With this information, you can produce the five verification conditions for proving the *partial* correctness of the program (one for showing that the input condition implies the invariant, one for showing that the invariant and the negation of the loop condition implies the output condition, three for showing that the invariant is preserved for each of the three possible execution paths in the loop body). Please note that the program term $\lfloor x \rfloor$ is written in PVS as `floor`$(x/2)$. Make sure that your definitions type-check correctly by using the PVS command `tcp` (no type checking condition should remain unproved).

Your task is now to verifiy these five conditions in PVS in the style of the proof of the "linear search" algorithm presented in class. For this purpose, write a PVS theory of the following structure

```
binarysearch: THEORY
BEGIN
  IMPORTING arrays[int]

  // program variables
  a, olda: arr
  x, oldx: int
  low, high, mid: int
  r: int

  // quantified variables
  i, j: VAR nat
  ...
END binarysearch
```

where the PVS file for theory `arrays` is posted on the Web site.

Define three predicates `Input`, `Output`, and `Invariant`, where (as shown in class) `Invariant` should be parameterized over the program variables. Then define five formulas `A`, `B1`, `B2`, `B3`, `C` describing the five verification conditions and prove these.

Please note that the program variables denoting indices are declared as `int` (r and *high* might become negative) such that all range restrictions on these variables have to be explicitly described. The quantified variables however may be declared as `nat` such that non-negativeness has not to be stated explicitly.

The following hints describe how the five conditions can be proved. The descriptions assume that you have expanded all predicate definitions.

A is a simple quantifier proof that can be performed using `split`, `flatten`, `skolem!`, `assert`.

B1 is even simpler than A.

B2 and B3 The proofs of both formulas are very similar. All but one branches of the proof run through automatically, but this one branch requires your attention. It needs for its proof an additional lemma

$$(\forall i : 0 \leq i < \mathit{length}(a) - 1 : a[i] \leq a[i+1]) \Rightarrow$$
$$(\forall i, j : 0 \leq i \leq j < \mathit{length}(a) : a[i] \leq a[j]).$$

Formulate this lemma as an additional formula `L` (which you need not prove) and introduce it into the proof by the use of the `lemma` command. Then apply the `split` command on this formula, which gives you one branch which is trivially fulfilled (because the antecedent of the lemma is satisfied), and another branch, where the consequent of the lemma is stated which you can use in the remaining proof. Apart from that, the proof only requires `split`, `flatten`, `skolem!`, `assert`, `inst`.

C requires a case distinction (command `case`) on r according to the two possible outcomes of the search. Otherwise it is a straight-forward quantifier proof which requires `split`, `flatten`, `skolem!`, `assert`, `inst` only.

I suggest that you first prove `A`, `B1`, and `C` and then `B2` and `B3`. All proofs have been checked, thus there is no reason that they should not work for you. However, if some proof does not work out completely, just submit the partial proof as your exercise result.

Bonus (20%): Also prove `L` using induction on i in the consequent formula.