# **Assignment 3**

Ruofei HUANG(z5141448) Anqi ZHU(z5141541) May 30, 2018

# 1 Task 1

# 1.1 Some useful symbol definitions about manipulating words

A word, as mentioned in the assignment specification, is defined as a finite sequence of letters over  $L = \{'a', ..., 'z'\}$ . The specification already defines the relationship of ' $\leq$ ' between a word v and a word w (this means v is a prefix of w or w itself if  $v \leq w$ ).

So we would like to further this definition and define a relationship of '<' when v is a proper prefix of w which cannot be w itself if v < w. That means:

$$v < w \Leftrightarrow v \le w \land v \ne w$$

We also would like to define a symbol |w| that represents the length of a word w. We formally define this by:

$$|w| = \begin{cases} 0 & \text{if } w = \epsilon \\ 1 + |w'| & \text{else} \end{cases}$$

where  $\exists l \in L \{w = w' \ l\}$ 

# 1.2 Syntactic(Abstract) Data Type Dict

Inspired by the program sketch and the assignment statement, we could describe the syntactic data type Dict as below (the encapsulated state would be a dictionary word set W).

$$Dict = (W = \phi,$$

$$\begin{pmatrix} \mathbf{proc} \ addword^{Dict}(\mathbf{word} \ w) \cdot b, W : [\mathtt{TRUE}, b = b_0 \wedge W = W_0 \cup \{w\}] \\ \mathbf{func} \ checkword^{Dict}(\mathbf{word} \ w) : \mathbb{B} \cdot \\ \mathbf{var} \ b \cdot b, W : [\mathtt{TRUE}, b = (w \in W) \wedge W = W_0]; \ \mathbf{return} \ b \\ \mathbf{proc} \ delword^{Dict}(\mathbf{word} \ w) \cdot b, W : [w \in W, b = b_0 \wedge W = W_0 \setminus \{w\}] \end{pmatrix})$$

# 2 Task 2

# 2.1 Data Type Refinement

Now we would like to refine Dict to a second data type DictA where we replace W with a trie t. We would also like to define the domain of a t as  $\mathbf{dom}(t)$ . We shall use this definition later in our refinement.

#### 2.1.1 Inductive Relation Predicate

The correspondence between the two state space W and t is captured by the inductively defined predicate as follows:

$$r = (W = \{w \in \mathbf{dom}(t) | t(w) = 1\})$$

We can translate this into a relation function that transfers a concrete state space t to an abstract state space W. The function is as follows:

$$f(t) = \{ w \in \mathbf{dom}(t) | t(w) = 1 \}$$

With that in mind, we can propose the initialisation predicate and corresponding operations of DictA.

#### 2.1.2 Initialisation Predicate

We would like to define the initialisation predicate of DictA as follows:

$$init^{DictA} = (t := \{\epsilon \mapsto 0\})$$

## 2.1.3 Operations

We would like to define the operations of DictA as follows:

```
\begin{aligned} &\mathbf{proc}\ addword^{DictA}(\mathbf{word}\ w) \cdot b, t: \\ &[\mathsf{TRUE}, b = b_0 \wedge t = t_0 \backslash \{w \mapsto 0\} \cup \{w \mapsto 1\} \cup \{w' < w \wedge w' \notin \mathbf{dom}(t) | w' \mapsto 0\}] \\ &\mathbf{func}\ checkword^{DictA}(\mathbf{word}\ w) : \mathbb{B} \cdot \mathbf{var}\ b \cdot b, t: \\ &[\mathsf{TRUE}, t = t_0 \wedge b = (\forall w' \leq w \, (w' \in \mathbf{dom}(t)) \wedge t(w) = 1)];\ \mathbf{return}\ b \\ &\mathbf{proc}\ delword^{DictA}(\mathbf{word}\ w) \cdot b, t: \\ &[\mathsf{TRUE}, b = b_0 \wedge (w \notin \mathbf{dom}(t) \vee t = (t : w \mapsto 0))] \end{aligned}
```

#### 2.2 Proof of Refinement

Now we would like to start proving the refinements of the initialization and each operation from t to W.

### 2.2.1 Refinement proof for init

We start from proving the refinement between  $init^{DictA}$  and  $init^{Dict}$ .

$$init^{DictA} \Rightarrow init^{Dict}[f(t)/W]$$
  
 $\Leftrightarrow \quad \langle \text{Definition of } init^{DictA} \text{ and } init^{Dict} \rangle$   
 $\forall w \in \mathbf{dom}(t) (t(w) = 0) \Rightarrow W = \phi$ 

Then we move on to prove the refinement of defined operations of Dict and DictA. We don't need to prove the validity of the condition  $(3_f)$  for any operations since their preconditions are always TRUE. We only need to check the validity of the condition  $(4_f)$  in all three operations.

### **2.2.2 Refinement proof for** addword

$$pre_{addword^{Dict}}[f(t_0)/W] \wedge post_{addword^{DictA}}$$

$$\Leftrightarrow \quad \langle \text{Definition of } addword^{Dict} \text{ and } addword^{DictA} \rangle$$

$$\text{TRUE}[f(t_0)/W] \wedge b = b_0 \wedge t = t_0 \backslash \{w \mapsto 0\} \cup \{w \mapsto 1\} \cup \{w' < w \wedge w' \notin \mathbf{dom}(t) | w' \mapsto 0\}$$

$$\Rightarrow \quad \langle \text{Definition of } f \rangle$$

$$f(t) = f(t_0 \backslash \{w \mapsto 0\} \cup \{w \mapsto 1\} \cup \{w' < w \wedge w' \notin \mathbf{dom}(t) | w' \mapsto 0\})$$

$$\Leftrightarrow \quad \langle \text{Logic, only } w \text{ maps to } 1 \text{ in } t(\text{only } w \text{ is newly added into } W) \rangle$$

$$f(t) = f(t_0) \cup \{w\}$$

$$\Leftrightarrow \quad \langle \text{Definition of } addword^{Dict} \text{ and } addword^{DictA} \rangle$$

$$post_{addword^{Dict}}[f(t_0),f(t)/W_0]$$

#### **2.2.3** Refinement proof for *checkword*

$$pre_{checkword^{Dict}}[f^{(t_0)}/W] \wedge post_{checkword^{DictA}}$$

$$\Leftrightarrow \quad \langle \text{Definition of } checkword^{DictA} \text{ and } checkword^{Dict} \rangle$$

$$\text{TRUE}[f^{(t_0)}/W] \wedge t = t_0 \wedge b = (\forall w' \leq w \, (w' \in \mathbf{dom}(t)) \wedge t(w) = 1)$$

$$\Rightarrow \quad \langle \text{Definition of } f \rangle$$

$$b = (w \in f(t) \wedge t(w) = 1) \wedge t = t_0$$

$$\Leftrightarrow \quad \langle \text{Definition of } checkword^{DictA} \text{ and } checkword^{Dict} \rangle$$

$$post_{checkword^{Dict}}[f^{(t_0)},f^{(t)}/W_0,W]$$

### 2.2.4 Refinement proof for delword

$$pre_{delword^{Dict}}[f^{(t_0)}/W] \wedge post_{delword^{DictA}}$$

$$\Leftrightarrow \quad \langle \text{Definition of } delword^{DictA} \text{ and } delword^{Dict} \rangle$$

$$w \in f(t_0) \wedge b = b_0 \wedge (w \notin \mathbf{dom}(t) \vee t = (t : w \mapsto 0))$$

$$\Rightarrow \quad \langle \text{Definition of } f \rangle$$

$$f(t) = f(t_0) \backslash \{w\}$$

$$\Leftrightarrow \quad \langle \text{Definition of } delword^{DictA} \text{ and } delword^{Dict} \rangle$$

$$post_{delword^{Dict}}[f^{(t_0),f(t)}/W_0,W]$$

# 3 Task 3

#### 3.1 Pre-defined function calls

Before refining the operations in DictA into toy language, we would like to first define some useful function calls that helps building our later refinement more close to the real c program constructions.

#### **3.1.1** *POPWORD*

The semantic function POPWORD returns a substring w' of the word w starting from the first letter to the index'th letter of w. The definition of the function and corresponding function call is as below:

#### **Definition of** *POPWORD*

$$POPWORD(w, i) = w'$$
 where  $w' \le w \land |w'| = i$ 

Definition of the function call popWord

func 
$$popWord(\mathbf{value}\ w, \mathbf{value}\ i) : \mathbf{word} \cdot \mathbf{var}\ w' \cdot \mathbf{u}w' : [i \leq |w|, w' \leq w \wedge |w'| = i]; \mathbf{u}_{(P1)}\mathbf{return}\ w'$$

Refinement of the procedure in popWord

$$(P1) \sqsubseteq \langle |w[0..i-1]| = i \wedge w[0..i-1] \leq w \rangle$$
$$w' := w[0..i-1]$$

This function helps us to build our refinement as close to the trie construction in c programs as possible. (In c programs, words are splitted into prefixes in increasing length and checked in the trie recursively.)

#### **3.1.2** doAddword

Using the previous definition of POPWORD, we would like to develop a corresponding function call named  $doAddword(\mathbf{var}\ w, \mathbf{var}\ index)$  that allows using a variable i to locate which prefix of  $w(\mathbf{or}\ \text{say}\ \text{which}\ \text{sub-trie}\ \text{of}\ t)$  the procedure currently is checking at during the entire word-adding operation.

All the prefixes of w should exist in t eventually, so if the current prefix POPWORD(w,i) does not exist, doAddword will add the new prefix into t and continue searching for the next prefix POPWORD(w,i+1). Considering this and the fact that t always initializes with  $\epsilon$  included, it is guaranteed that every doAddword operation at i level will already have all prefixes of w with length no longer than i exist in t, which satisfies the precondition of the function itself. The definition of the function and the refinement of its linking procedure is as follows.

#### Definition of doAddword

#### Refinement of the procedure in doAddword

```
(A2) \square
                                                                                                \langle if \rangle
                                                                          if index < |w|
                                                                          then t: [index < |w| \land pre(A2), post(A2)] \rfloor_{(A3-1)}
                                                                          else_{\perp}t : [index \geq |w| \land pre(A2), post(A2)]_{(A3-2)}
                                                                          fi
                   (A3-1) \square
                                                                                               \langle \text{seq} \rangle
                                                                        \begin{split} & \mathbf{L}t: \left[ \begin{array}{l} index \geq |w| \wedge pre(A2), \\ index \geq |w| \wedge pre(A2) \wedge POPWORD(w, index + 1) \in \mathbf{dom}(t) \end{array} \right] \mathbf{L}(A3-3) \\ & \mathbf{L}t: \left[ \begin{array}{l} index \geq |w| \wedge pre(A2) \wedge POPWORD(w, index + 1) \in \mathbf{dom}(t), \\ post(A2) \end{array} \right] \mathbf{L}(A3-4) \end{split}
                  (A3-3) \sqsubseteq
                                                                          if (POPWORD(w, index + 1) \in \mathbf{dom}(t))
                                                                        then \[ \] then \[ \
                                                                                          \langle POPWORD(w, index + 1) \in \mathbf{dom}(t) \Rightarrow post(A3 - 3) = \text{TRUE} \rangle
(A3 - 3 - 1) \sqsubseteq
                                                                          skip;
(A3 - 3 - 2) \sqsubseteq \langle Ass \rangle
                                                                        t := t \cup \{popWord(w, index) \mapsto 0\}
                   (A3-4) \sqsubseteq \langle \text{proc} \rangle
                                                                          doAddword(w, index + 1);
                  (A3-2) \sqsubseteq \langle index \geq |w| \land index \leq |w| \Rightarrow index = |w| \text{ and definition of } post(A2) \rangle
                                                                          t := t_0 \setminus \{w \mapsto 0\} \cup \{w \mapsto 1\};
```

We gather the code for the body of doAddword

```
\begin{array}{l} \textbf{if} \quad index < |w| \\ \textbf{then} \\ \quad \textbf{if} \quad POPWORD(w,index+1) \in \textbf{dom}(t) \\ \textbf{then} \quad skip; \\ \textbf{else} \\ \quad \quad t := t \cup \{popWord(w,index) \mapsto 0\} \\ \quad \textbf{fi} \\ \quad \quad doAddword(w,index+1); \\ \textbf{else} \\ \quad \quad t := t_0 \backslash \{w \mapsto 0\} \cup \{w \mapsto 1\}; \\ \textbf{fi} \end{array}
```

#### **3.1.3** doCheckword

Similarly to doAddword, we would also develop a corresponding function call named  $doCheckword(\mathbf{var}\ w, \mathbf{var}\ index)$  which will search the complete word w in t based on the precondition that all prefixes of w with length no longer than index are guaranteed to exist in t. The definition of the function and the refinement of its linking procedure is as follows.

#### Definition of doCheckword

```
\begin{aligned} & \mathbf{func} \ doCheckword(\mathbf{value} \ w, \mathbf{value} \ index : \mathbb{N}) : \mathbb{B} \cdot \\ & \mathbf{var} \ b \cdot \mathbf{L}b : \begin{bmatrix} 0 \leq index \leq |w| \wedge \\ \forall w' \leq w \wedge |w'| \leq index \, (w' \in \mathbf{dom}(t)) \, , \\ b = (\forall w' \leq w \, (w' \in \mathbf{dom}(t)) \wedge t(w) = 1) \end{bmatrix} \mathbf{l}_{(C2)}; \mathbf{return} \ b \end{aligned}
```

### Refinement of the procedure in doCheckword

```
(C2) \sqsubseteq \langle \text{if} \rangle
\text{if } index < |w|
\text{then} \bot b : [index < |w| \land pre(C2), post(C2)] \bot_{(C3-1)}
\text{else} \bot b : [index \ge |w| \land pre(C2), post(C2)] \bot_{(C3-2)}
\text{fi}
(C3-1) \sqsubseteq \langle \text{if} \rangle
\text{if } (POPWORD(w, index + 1) \in \text{dom}(t))
\text{then } \bot b : \begin{bmatrix} pre(C3-1) \land POPWORD(w, index + 1) \in \text{dom}(t), \\ post(C2) \end{bmatrix} \bot_{(C3-1-1)}
\text{else } \bot b : \begin{bmatrix} pre(C3-1) \land POPWORD(w, index + 1) \notin \text{dom}(t), \\ post(C2) \end{bmatrix} \bot_{(C3-1-2)}
\text{fi}
(C3-1-1) \sqsubseteq \langle \text{ass, func} \rangle
b := doCheckword(w, index + 1);
(C3-1-2) \sqsubseteq \langle \text{Ass and } popWord(w, index + 1) \notin \text{dom}(t) \Rightarrow b = \text{FALSE} \rangle
b := \text{FALSE}
(C3-2) \sqsubseteq \langle index \ge |w| \land index \le |w| \Rightarrow index = |w| \text{ and definition of } post(C2) \rangle
b := (t(w) = 1)
```

We gather the code of doCheckword below:

```
\begin{aligned} &\textbf{if} \quad index < |w| \\ &\textbf{then} \\ &\textbf{if} \quad POPWORD(w,index+1) \in \textbf{dom}(t) \\ &\textbf{then} \quad skip; \\ &\textbf{else return} \quad doCheckword(w,index+1); \\ &\textbf{fi} \\ &\textbf{else} \quad b := (t(w) = 1); \\ &\textbf{fi} \end{aligned}
```

#### **3.1.4** doDelword

Similarly to doCheckword, we would also develop a corresponding function call named  $doDelword(\mathbf{var}\ w, \mathbf{var}\ index)$  which will set the result of t(w) as 0 based on the precondition that all prefixes of w with length no longer than index are guaranteed to exist in t. The definition of the function and the refinement of its linking procedure is as follows.

#### Definition of doDelword

# Refinement of the procedure in doDelword

```
(D2) \sqsubseteq
                              \langle if \rangle
                              if index < |w|
                              then t: [index < |w| \land pre(D2), post(D2)] \rfloor_{(D3-1)}
                              else_{L}t : [index \ge |w| \land pre(D2), post(D2)]_{(C3-2)}
       (D3-1) \sqsubseteq
                                      \langle if \rangle
                              if (POPWORD(w, index + 1) \in \mathbf{dom}(t))
                             \begin{aligned} &\textbf{then} \ \llcorner t: \left[\begin{array}{c} pre(D3-1) \land POPWORD(w,index+1) \in \textbf{dom}(t), \\ post(D2) \end{array}\right] \lrcorner (D3-1-1) \end{aligned} &\textbf{else} \ \llcorner t: \left[\begin{array}{c} pre(D3-1) \land POPWORD(w,index+1) \notin \textbf{dom}(t), \\ post(D2) \end{array}\right] \lrcorner (D3-1-2) \end{aligned}
                                      (ass, func)
(D3 - 1 - 1) \sqsubseteq
                              doDelword(w, index + 1);
                                    (Ass and POPWORD(w, index + 1) \notin \mathbf{dom}(t) \Rightarrow post(D2) = \text{TRUE})
(D3 - 1 - 2) \sqsubseteq
                             skip;
       (D3-2) \sqsubseteq \qquad \langle index \geq |w| \wedge index \leq |w| \Rightarrow index = |w| \text{ and definition of } post(D2) \rangle
                             t := t : w \mapsto 0:
```

We gather the code of doDelword below:

```
\begin{aligned} &\textbf{if} \quad index < |w| \\ &\textbf{then} \\ &\textbf{if} \quad POPWORD(w,index+1) \in \mathbf{dom}(t) \\ &\textbf{then} \ doDelword(w,index+1); \\ &\textbf{else} \ skip; \\ &\textbf{fi} \\ &\textbf{else} \ t := t : w \mapsto 0; \\ &\textbf{fi} \end{aligned}
```

#### 3.2 Refinement of *init*

From the spec we have:

$$\mathbf{dom}(t) = \{\epsilon\} \land f(t) = \phi$$

$$\sqsubseteq \quad \langle \text{ass} \rangle$$

$$t := \{\epsilon \mapsto 0\}$$

### 3.3 Refinement of addword

From the spec<sup>1</sup> we have:

#### **3.4 Refinement of** *checkword*

From the spec we have:

```
func checkword^{DictA}(\mathbf{value}\ w): \mathbb{B} \cdot 
\mathbf{var}\ b \cdot \llcorner b, t : [\mathsf{TRUE}, b = (\forall w' \leq w\ (w' \in \mathbf{dom}(t)) \land t(w) = 1)]; \lrcorner_{(C1)}\ \mathbf{return}\ b
(C1) \sqsubseteq \qquad \langle \mathsf{c-frame} \rangle 
b : [\mathsf{TRUE}, b = (\forall w' \leq w\ (w' \in \mathbf{dom}(t)) \land t(w) = 1)];
\sqsubseteq \qquad \langle \mathsf{proc}, \ 0 \leq |w| \ \mathsf{and}\ \epsilon \in \mathbf{dom}(t) \ \mathsf{since}\ init^{DictA} \rangle 
b := doCheckword(w, 0);
```

### **3.5** Refinement of delword

From the spec we have:

<sup>&</sup>lt;sup>1</sup>Definition of this is in the Assignment 3 requirements of cs2111.

# 3.6 Task 4

Now we would translate our data refinement to C functions to match the prototypes given in dict.h. There are some significant differences between the structure of a trie used in c functions and in our data refinement (for example, the tries in c functions can be operated as nodes but they are operated as sets in our data refinement). However, it should be noted that the logic used on both sides is the same (they both use recursive call to derive the final result).

```
#include "dict.h"
 1
    #include <stdio.h>
 3
    #include <stdlib.h>
    void newdict(Dict *dp){
         // malloc the space of root node.
 5
 6
        *dp = malloc(sizeof(struct \_tnode\_));
 7
 8
 9
    void doAddword(const Dict r, const word w, int index) {
10
        if (w[index]!= '\0') {
11
             // the word is not ended
12
            if (r->cvec[w[index+1]-'a']!=NULL) {
13
                 /* POPWORD(w,index+1) \setminus in dom(t) */
14
                 // skip
15
             }
            else{}
16
                 //t := (POPWORD(w, index+1) -> 0)
17
18
                 \operatorname{newdict}(\&(r->\operatorname{cvec}[w[\operatorname{index}+1]-'a']));
19
20
             // recursive call
21
            doAddword(r->cvec[w[index+1]-'a'],w, index +1);
22
        }
23
        else{}
             //index = |w|
24
            //t:=t_-\theta\setminus\{w->\theta\}\ U\{w->1\}
25
            r->eow = TRUE;
26
        }
27
28
    }
29
30
    bool doCheckword(const Dict r, const word w, int index){
31
        if (w[index]!= '\0') {
32
             // the word is not ended
            if (r->cvec[w[index+1]-'a']!=NULL) {
33
34
                 /* POPWORD(w,index+1) \setminus in dom(t) */
35
                 // skip
             }
36
```

```
37
            else{
                // t := (POPWORD(w, index+1) -> 0)
38
39
                // there is not exist the word in this dict
                return FALSE;
40
41
42
            // recursive call
            return doCheckword(r->cvec[w[index+1]-'a'],w, index +1);
43
        }
44
        \mathbf{else} \{
45
46
            //index = |w|
47
            // return b := (t(w) = 1)
48
            return r->eow;
49
        }
50
    }
51
52
    void doDelword(const Dict r, const word w, int index) {
53
        if (w[index]!= '\0') {
54
            // the word is not ended
55
            if (r->cvec[w[index+1]-'a']!=NULL) {
                /* POPWORD(w,index+1) \setminus in dom(t) */
56
57
                 // recursive call
                doDelword(r->cvec[w[index+1]-'a'],w, index+1);
58
59
            \mathbf{else} \{
60
                // t := (POPWORD(w, index+1) \rightarrow 0)
61
62
                // there is not exist the word in this dict
                // nothing to delete
63
64
                // skip;
65
                return;
66
67
        }
68
        else{
69
            //index = |w|
            // return t:=t:w->0
70
            r \rightarrow eow = FALSE;
71
72
        }
73
    }
74
    void addword (const Dict r, const word w){
75
        doAddword(r, w,0);
76
77
   }
78
   bool checkword (const Dict r, const word w){
        return doCheckword(r, w,0);
79
80
    }
```