

Formal Specification Examples

Rui Carlos A. Gonçalves (41031)

January 6, 2007

Contents

1. Problem 2 - Modelling with Sequences	1
1.1. Pretty-print	1
1.2. Source code	2
2. Problem 3 - Modelling with Sets	4
2.1. Pretty-print	5
2.2. Source code	7
3. Problem 6 - Modelling with Mappings	10
3.1. Pretty-print	11
3.2. Source code	12
4. Problem 7 - Multisets “are” Mappings	14
4.1. Pretty-print	15
4.2. Source code	16
5. Problem 8 - Hash tables “are” Mappings	17
5.1. Pretty-print	17
5.2. Source code	19
6. Problem 10 - A Naive Model of the WWW	20
6.1. Pretty-print	20
6.2. Source code	23
7. Problem 11 - Modelling a Memory Cache	25
7.1. Pretty-print	26
7.2. Source code	29

1. Problem 2 - Modelling with Sequences

1. The squadrons of a large military contingent can be in two possible states: operational or resting. Once a squadron is in operation for some time limit, it must return to the military base and change to the resting state, being replaced by the squadron which has the longest rest.

Use `seq` of modelling to complete the ... in

```
Contingent :: operational: ...
             resting: ...
```

and specify the operation of squadron replacement outlined above.

2. Use sequence comprehension to model the following requirements from a mobile phone manufacturer:

(...) For each list of calls stored in the mobile phone (eg. numbers dialed, SMS messages, lost calls), the store operation should work in a way such that (a) the more recently a call is made the more accessible it should be; (b) no number appears twice in a list; (c) only the last 10 entries in each list are stored.

3. Transliterate the following (polymorphic) function definition to VDM-SL syntax

```
filter p [] = []
filter p (a:l) | p a      = a : (filter p l)
                | otherwise = (filter p l)
```

and check its behaviour in the toolbox.

1.1. Pretty-print

types

- 1.0 $Contingent :: operational : \mathbb{Z}^*$
- .1 $resting : \mathbb{Z}^*$
- .2 $inv\ c \triangleq elems\ (c.operational) \cap elems\ (c.resting) = \{\}$

functions

- 2.0 $replace : Contingent \rightarrow Contingent$
- .1 $replace\ (x) \triangleq$
- .2 $mk\ Contingent\ (tl\ (x.operational) \curvearrowright [hd\ (x.resting)],$
- .3 $tl\ (x.resting \curvearrowright [hd\ (x.operational)]))$
- .4 $pre\ x.operational \neq [] \wedge x.resting \neq [] ;$

```

3.0 addcall :  $\mathbb{Z} \times \mathbb{Z}^* \rightarrow \mathbb{Z}^*$ 
.1 addcall (n, l)  $\triangleq$ 
.2   if card (elems (l  $\curvearrowright$  [n])) = 10  $\vee$  len (l) < 10
.3   then [n]  $\curvearrowright$  [l(i) | i  $\in$  inds l  $\cdot$  l(i)  $\neq$  n]
.4   else [n]  $\curvearrowright$  [l(i) | i  $\in$  {1, ..., 9}]
.5 pre len (l)  $\leq$  10  $\wedge$  len (l) = card (elems (l)) ;

```

```

4.0 filter[@a] : (@a  $\rightarrow$   $\mathbb{B}$ )  $\times$  @a*  $\rightarrow$  @a*
.1 filter (p, l)  $\triangleq$ 
.2   cases l :
.3     []  $\rightarrow$  [],
.4     [h]  $\curvearrowright$  t  $\rightarrow$  if p (h)
.5       then [h]  $\curvearrowright$  filter[@a] (p, t)
.6       else filter[@a] (p, t)
.7   end;

```

```

5.0 lessSix :  $\mathbb{Z} \rightarrow \mathbb{B}$ 
.1 lessSix (x)  $\triangleq$ 
.2   (x < 6);

```

```

6.0 prefixA : char*  $\rightarrow \mathbb{B}$ 
.1 prefixA (c)  $\triangleq$ 
.2   (hd (c) = 'a');

```

values

```

7.0 c1 = mk-Contingent ([1, 2, 3], [4, 5, 6, 7, 8, 9, 0]);

```

```

8.0 c2 = mk-Contingent ([], [2, 3, 4, 5]);

```

```

9.0 calls = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

```

1.2. Source code

```

-----
-- Rui Carlos A. Goncalves
-- LMCC - MFP I
-- 2006/07
-- Problema 2
-----

```

```

types
Contingent::operational:seq of int
           resting      :seq of int
inv c==elems(c.operational) inter elems(c.resting)={};

functions
-- 1
replace:Contingent->Contingent
replace(x)==mk_Contingent(tl(x.operational)^[hd(x.resting)],
                          tl(x.resting^[hd(x.operational)]))
pre x.operational<>[] and x.resting<>[];

-- 2
addcall:int*seq of int->seq of int
addcall(n,l)==if card(elems(l^[n]))=10 or len(l)<10
               then [n]^[l(i)|i in set inds l & l(i)<n]
               else [n]^[l(i)|i in set {1,...,9}]
pre len(l)<=10 and len(l)=card(elems(l));

--3
filter[@a]:(@a->bool)*seq of @a->seq of @a
filter(p,l)==cases l:
             []->[],
             [h]^t->if p(h)
                    then [h]^filter[@a](p,t)
                    else filter[@a](p,t)
             end;

lessSix:int->bool
lessSix(x)==(x<6);

prefixA:seq of char->bool
prefixA(c)==(hd(c)='a');

values
c1=mk_Contingent([1,2,3],[4,5,6,7,8,9,0]);
c2=mk_Contingent([], [2,3,4,5]);

calls=[1,2,3,4,5,6,7,8,9,10];

```

2. Problem 3 - Modelling with Sets

To improve the operation of fire-brigades fighting forest fires in the hot season, the National Service for Civil Protection is developing a mobile station-based geographical information system which keeps track of rural area accessibility.

In this exercise we are concerned with the design of the central database component which records road maps (distances, road alternatives and so on) - a component which we are going to model with sets.

We start from a very simple model (to be improved at a later stage): a road-map is a directed graph whose arcs $A \rightarrow B$ record the *you can drive from A to B* relation:

```
Map = set of Path;
Path :: From: Location
       To: Location;
Location = token;
```

1. Specify functions in VDM-SL to obtain:
 1. The symmetric of a map (ie. the map of the “all ways back”);
 2. The symmetric closure of a map (ie. the map and all its ways back);
 3. The composition of two maps (ie. the map whose paths are obtained by arriving at a location accessible in the first map and proceeding from there to another location accessible in the second map);
 4. The reflection of a map (set of all trivial paths between a location in the map and itself);
 5. The transitive closure of a map (set of all direct or composed paths in a map).
2. The combined symmetric, transitive and reflexive closure of a map is a map which records an equivalence relation on locations. What is the informal meaning of this equivalence?
3. Upgrade your Map model by adding further information to each path, namely the **name** of the road, the **distance** involved and its (average) **speed**:

```
UMap = set of UPath;
UPath :: From: Location
        Info: PathInfo
        To: Location;
PathInfo :: name: token
           distance: real
           speed: real;
Location = token;
```

Scale up the function definitions above from Map to UMap. Clearly, you have to pay attention to the handling of PathInfos, by eg. summing up distances, computing average speeds and concatenating road names. (**Hint**: recall the notion of a *weighted graph*.)

2.1. Pretty-print

types

10.0 $Location = \text{token};$

11.0 $Map = \text{Path-set};$

12.0 $Path :: From : Location$
.1 $To : Location;$

13.0 $UMap = UPath\text{-set};$

14.0 $UPath :: From : Location$
.1 $Info : PathInfo$
.2 $To : Location;$

15.0 $PathInfo :: name : \text{token}$
.1 $distance : \mathbb{R}$
.2 $speed : \mathbb{R}$

functions

16.0 $symmetric : Map \rightarrow Map$
.1 $symmetric(m) \triangleq$
.2 $\{\text{mk-Path}(x.To, x.From) \mid x \in m\};$

17.0 $symClosure : Map \rightarrow Map$
.1 $symClosure(m) \triangleq$
.2 $symmetric(m) \cup m;$

18.0 $composition : Map \times Map \rightarrow Map$
.1 $composition(m1, m2) \triangleq$
.2 $\{\text{mk-Path}(x, z) \mid$
.3 $\text{mk-Path}(x, y) \in m1,$
.4 $\text{mk-Path}(y, z) \in m2\};$

19.0 $reflection : Map \rightarrow Map$
.1 $reflection(m) \triangleq$
.2 $\{\text{mk-Path}(x, x) \mid \text{mk-Path}(x, y) \in m\} \cup$
.3 $\{\text{mk-Path}(y, y) \mid \text{mk-Path}(x, y) \in m\} \cup$
.4 $m;$

```

20.0 transClosure : Map → Map
.1 transClosure (m)  $\triangleq$ 
.2   if m = composition (m, m) ∪ m
.3   then m
.4   else transClosure (composition (m, m) ∪ m);

21.0 usymmetric : UMap → UMap
.1 usymmetric (m)  $\triangleq$ 
.2   {mk-UPath (x.To, x.Info, x.From) | x ∈ m};

22.0 usymClosure : UMap → UMap
.1 usymClosure (m)  $\triangleq$ 
.2   usymmetric (m) ∪ m;

23.0 ucomposition : UMap × UMap → UMap
.1 ucomposition (m1, m2)  $\triangleq$ 
.2   {mk-UPath (x, mk-PathInfo (mk-token ({inf1, inf2}),
.3     d1 + d2,
.4     (s1 + s2)/2),
.5     z) |
.6     mk-UPath (x, mk-PathInfo (inf1, d1, s1), y) ∈ m1,
.7     mk-UPath (y, mk-PathInfo (inf2, d2, s2), z) ∈ m2};

24.0 ureflection : UMap → UMap
.1 ureflection (m)  $\triangleq$ 
.2   {mk-UPath (x, mk-PathInfo (mk-token ("NULL"), 0, 0), x) |
.3     mk-UPath (x, inf, y) ∈ m} ∪
.4   {mk-UPath (y, mk-PathInfo (mk-token ("NULL"), 0, 0), y) |
.5     mk-UPath (x, inf, y) ∈ m} ∪
.6   m;

25.0 utransClosure : UMap → UMap
.1 utransClosure (m)  $\triangleq$ 
.2   if m = ucomposition (m, m) ∪ m
.3   then m
.4   else utransClosure (ucomposition (m, m) ∪ m)

```

values

```

26.0 l1 : token = mk-token ('A');

```

```

27.0  l2 : token = mk-token ('B');

28.0  l3 : token = mk-token ('C');

29.0  l4 : token = mk-token ('D');

30.0  l5 : token = mk-token ('E');

31.0  l6 : token = mk-token ('F');

32.0  m1 : Map = {mk-Path (l1, l2), mk-Path (l3, l2), mk-Path (l2, l4)};

33.0  m2 : Map = {mk-Path (l2, l5), mk-Path (l4, l6)};

34.0  um1 : UMap = {mk-UPath (l1, mk-PathInfo (l1, 2, 3), l2),
.1      mk-UPath (l3, mk-PathInfo (l3, 3, 4), l2),
.2      mk-UPath (l2, mk-PathInfo (l2, 1, 4), l4)}

```

2.2. Source code

```

-----
-- Rui Carlos A. Goncalves
-- LMCC - MFP I
-- 2006/07
-- Problema 3
-----

types
Location = token;

Map = set of Path;
Path :: From : Location
      To : Location;

UMap = set of UPath;
UPath :: From: Location
        Info: PathInfo
        To: Location;
PathInfo :: name: token
           distance: real
           speed: real;

```



```

functions
-- 1
-- 1.1
symmetric:Map->Map
symmetric(m)={mk_Path(x.To,x.From)|x in set m};

-- 1.2
symClosure:Map->Map
symClosure(m)={symmetric(m) union m};

-- 1.3
composition:Map*Map->Map
composition(m1,m2)={mk_Path(x,z)
                    |mk_Path(x,y) in set m1
                    ,mk_Path(y,z) in set m2};

-- 1.4
reflection:Map->Map
reflection(m)={mk_Path(x,x)|mk_Path(x,y) in set m}
              union {mk_Path(y,y)|mk_Path(x,y) in set m}
              union m;

-- 1.5
transClosure:Map->Map
transClosure(m)={if m=composition(m,m) union m
                 then m
                 else transClosure(composition(m,m) union m);

-- 2
-- Se no grafo original e possivel chegar de um nodo a outro, no final vai
-- existir uma ligacao directa entre eles nos dois sentidos.

-- 3
-- 3.1
usymmetric:UMap->UMap
usymmetric(m)={mk_UPath(x.To,x.Info,x.From)|x in set m};

-- 3.2
usymClosure:UMap->UMap
usymClosure(m)={usymmetric(m) union m};

-- 3.3
ucomposition:UMap*UMap->UMap
ucomposition(m1,m2)={mk_UPath(x,mk_PathInfo(mk_token({inf1,inf2})
                                             ,d1+d2
                                             ,(s1+s2)/2)
                                             ,z)
                    |mk_UPath(x,mk_PathInfo(inf1,d1,s1),y) in set m1
                    ,mk_UPath(y,mk_PathInfo(inf2,d2,s2),z) in set m2};

-- 3.4
ureflection:UMap->UMap

```

```

ureflection(m)=={mk_UPath(x,mk_PathInfo(mk_token("NULL"),0,0),x)
                |mk_UPath(x,inf,y) in set m}
union {mk_UPath(y,mk_PathInfo(mk_token("NULL"),0,0),y)
      |mk_UPath(x,inf,y) in set m}
union m;

-- 3.5
utransClosure:UMap->UMap
utransClosure(m)==if m=ucomposition(m,m) union m
                  then m
                  else utransClosure(ucomposition(m,m) union m);

values
l1:token=mk_token('A');
l2:token=mk_token('B');
l3:token=mk_token('C');
l4:token=mk_token('D');
l5:token=mk_token('E');
l6:token=mk_token('F');
m1:Map={mk_Path(11,12),mk_Path(13,12),mk_Path(12,14)};
m2:Map={mk_Path(12,15),mk_Path(14,16)};

um1:UMap={mk_UPath(11,mk_PathInfo(11,2,3),12),
          mk_UPath(13,mk_PathInfo(13,3,4),12),
          mk_UPath(12,mk_PathInfo(12,1,4),14)};

```

3. Problem 6 - Modelling with Mappings

1. Complete the “.....” in the VDM-SL model below of a toy *bank account management system* (BAMS) in which accounts (**Account**) are uniquely identified by account identifiers (**AccId**). Every account records both a set of account holders (**set of AccHolder**) and the account’s current balance (**Amount**):

```

types

BAMS = map AccId to Account;
Account :: H: set of AccHolder
          B: Amount;

AccId    = seq of char;
AccHolder = seq of char;
Amount = int;

functions

Init : () -> BAMS
Init () == {|->};

OpenAccount : BAMS * AccId * set of AccHolder * Amount -> BAMS
OpenAccount(bams,n,h,m) ==
    bams union .....
pre not n in set dom bams;

AddAccHolders: BAMS * AccId * set of AccHolder -> BAMS
AddAccHolders(bams,n,h) ==
    bams ++ { n |-> let a = bams(n)
                in mk_Account(a.H union h,a.B) }
pre .....;

Credit: BAMS * AccId * Amount -> BAMS
Credit(bams,n,m) == .....
pre n in set dom bams;

Withdraw: BAMS * AccId * Amount -> BAMS
Withdraw(bams,n,m) ==
    bams ++ { n |-> let a = bams(n)
                    in mk_Account(a.H, a.B - m) }
pre .....;

GoodCostumers: BAMS * Amount -> set of AccId
GoodCostumers(bams,m) ==
    --- should select the accounts whose balance is greater than m

```

2. Attach an invariant to BAMS ensuring:
 1. that every account has at least one account holder;
 2. that the balance of every account is positive.

3.1. Pretty-print

types

- 35.0 $BAMS = AccId \xrightarrow{m} Account;$
- 36.0 $Account :: H : AccHolder\text{-set}$
 - .1 $B : Amount$
 - .2 $inv\ a \triangleq (a.H \neq \{\}) \wedge (a.B > 0);$
- 37.0 $AccId = char^*;$
- 38.0 $AccHolder = char^*;$
- 39.0 $Amount = \mathbb{Z}$

functions

- 40.0 $Init : () \rightarrow BAMS$
 - .1 $Init () \triangleq$
 - .2 $\{\mapsto\};$
- 41.0 $OpenAccount : BAMS \times AccId \times AccHolder\text{-set} \times Amount \rightarrow BAMS$
 - .1 $OpenAccount (bams, n, h, m) \triangleq$
 - .2 $bams \uplus \{n \mapsto mk\text{-Account} (h, m)\}$
 - .3 $pre\ \neg n \in \text{dom } bams ;$
- 42.0 $AddAccHolders : BAMS \times AccId \times AccHolder\text{-set} \rightarrow BAMS$
 - .1 $AddAccHolders (bams, n, h) \triangleq$
 - .2 $bams \dagger \{n \mapsto \text{let } a = bams (n) \text{ in}$
 - .3 $mk\text{-Account} (a.H \cup h, a.B)\}$
 - .4 $pre\ n \in \text{dom } bams ;$
- 43.0 $Credit : BAMS \times AccId \times Amount \rightarrow BAMS$
 - .1 $Credit (bams, n, m) \triangleq$
 - .2 $bams \dagger \{n \mapsto \text{let } a = bams (n) \text{ in}$
 - .3 $mk\text{-Account} (a.H, a.B + m)\}$
 - .4 $pre\ n \in \text{dom } bams ;$

44.0 $Withdraw : BAMS \times AccId \times Amount \rightarrow BAMS$

```
.1  $Withdraw(bams, n, m) \triangleq$   
.2  $bams \uparrow \{n \mapsto \text{let } a = bams(n) \text{ in}$   
.3  $\text{mk\_Account}(a.H, a.B - m)\}$   
.4  $\text{pre } n \in \text{dom } bams \wedge bams(n).B \geq m ;$ 
```

45.0 $GoodCostumers : BAMS \times Amount \rightarrow AccId\text{-set}$

```
.1  $GoodCostumers(bams, m) \triangleq$   
.2  $\{x \mid x \in \text{dom } bams \cdot bams(x).B > m\}$ 
```

3.2. Source code

```
-----  
-- Rui Carlos A. Goncalves  
-- LMCC - MFP I  
-- 2006/07  
-- Problema 6  
-----  
  
types  
BAMS = map AccId to Account;  
Account :: H: set of AccHolder  
          B: Amount  
inv a==(a.H<>{ }) and (a.B>0);  
  
AccId      = seq of char;  
AccHolder = seq of char;  
Amount = int;  
  
functions  
Init : () -> BAMS  
Init () == {|->};  
  
OpenAccount : BAMS * AccId * set of AccHolder * Amount -> BAMS  
OpenAccount(bams,n,h,m) ==  
    bams munion {n|->mk_Account(h,m)}  
pre not n in set dom bams;  
  
AddAccHolders: BAMS * AccId * set of AccHolder -> BAMS  
AddAccHolders(bams,n,h) ==  
    bams ++ { n |-> let a = bams(n)  
                  in mk_Account(a.H union h,a.B) }  
pre n in set dom bams;  
  
Credit: BAMS * AccId * Amount -> BAMS  
Credit(bams,n,m)==bams ++ {n |-> let a=bams(n)  
                              in mk_Account(a.H,a.B+m)}  
pre n in set dom bams;
```

```
Withdraw: BAMS * AccId * Amount -> BAMS
Withdraw(bams,n,m) ==
    bams ++ { n |-> let a = bams(n)
                in mk_Account(a.H, a.B - m) }
pre n in set dom bams and bams(n).B>=m;

GoodCostumers: BAMS * Amount -> set of AccId
GoodCostumers(bams,m)={x|x in set dom bams & bams(x).B>=m};
```

4. Problem 7 - Multisets “are” Mappings

A multiset (vulg. “bag”) is an “intermediate” notion between a set (**set of A**) and a sequence (**seq of A**). In a multiset, a given element a of A may occur more than once, its number of occurrences being called its *multiplicity*:

```
map @A to nat
```

Multisets are very common in formal modelling. For instance, an inventory

What	Quantity
chair	12
table	3
cupboard	1

is modelled by a multiset:

```
m = {"chair"|->12,"table"|->3,"cupboard"|->1}
```

Invoices are multisets, bank statements are multisets, etc.

The empty mapping $\{|->\}$ models the empty multiset and set-union generalizes to multiplicity addition, that is, given two multisets n and m , their union is defined as follows:

```
mseCup[@A] : map @A to nat * map @A to nat -> map @A to nat
mseCup(m,n) == m ++ n ++
               { a |-> m(a) + n(a) | a in set dom n inter dom m };
```

1. What is the outcome of $mseCup(m,m)$? Check it in the toolbox for m the multiset above. Check that $mseCup(m,m)$ has the same elements of m but that their multiplicity is two times larger. Generalize this by defining

```
mseExMul[@A] : map @A to nat * nat -> map @A to nat
```

the multiplication of a multiset by a numeric factor.

2. Define a function

```
mseDiff[@A] : map @A to nat * map @A to nat -> map @A to nat
```

which computes multiset *difference*.

3. Generalize multiset union to a finite set of multisets:

```
mseCUP[@A] : seq of map @A to nat -> map @A to nat
```

4.1. Pretty-print

functions

46.0 $mseCup[@A] : @A \xrightarrow{m} \mathbb{N} \times @A \xrightarrow{m} \mathbb{N} \rightarrow @A \xrightarrow{m} \mathbb{N}$

- .1 $mseCup(m, n) \triangleq$
- .2 $m \dagger n \dagger$
- .3 $\{a \mapsto m(a) + n(a) \mid a \in \text{dom } n \cap \text{dom } m\};$

47.0 $mseExMul[@a] : @a \xrightarrow{m} \mathbb{N} \times \mathbb{N} \rightarrow @a \xrightarrow{m} \mathbb{N}$

- .1 $mseExMul(m, n) \triangleq$
- .2 if $n = 1$
- .3 then m
- .4 else $mseCup[@a](m, mseExMul[@a](m, n - 1));$

48.0 $mseExMul2[@a] : @a \xrightarrow{m} \mathbb{N} \times \mathbb{N} \rightarrow @a \xrightarrow{m} \mathbb{N}$

- .1 $mseExMul2(m, n) \triangleq$
- .2 $\{a \mapsto m(a) \times n \mid a \in \text{dom } m\};$

49.0 $mseDiff[@a] : @a \xrightarrow{m} \mathbb{N} \times @a \xrightarrow{m} \mathbb{N} \rightarrow @a \xrightarrow{m} \mathbb{N}$

- .1 $mseDiff(m1, m2) \triangleq$
- .2 $\{a \mapsto (m1(a) - m2(a)) \mid a \in \text{dom } m1,$
- .3 $a \in \text{dom } m2 \cdot$
- .4 $m1(a) - m2(a) > 0\} \uplus$
- .5 $\{a \mapsto m1(a) \mid a \in \text{dom } m1 \cdot$
- .6 $\neg(a \in \text{dom } m2)\};$

50.0 $mseCUP[@a] : @a \xrightarrow{m} \mathbb{N}^* \rightarrow @a \xrightarrow{m} \mathbb{N}$

- .1 $mseCUP(s) \triangleq$
- .2 if $\text{len } s = 1$
- .3 then $\text{hd } s$
- .4 else $mseCup[@a](\text{hd } s, mseCUP[@a](\text{tl } s))$
- .5 pre $\text{len } s \geq 1$

values

51.0 $m1 = \{ "a" \mapsto 3, "b" \mapsto 3, "c" \mapsto 1, "d" \mapsto 1 \};$

52.0 $m2 = \{ "a" \mapsto 3, "c" \mapsto 4, "e" \mapsto 1, "g" \mapsto 5 \}$

4.2. Source code

```
-----  
-- Rui Carlos A. Goncalves  
-- LMCC - MFP I  
-- 2006/07  
-- Problema 7  
-----  
  
functions  
mseCup[@A] : map @A to nat * map @A to nat -> map @A to nat  
mseCup(m,n) == m ++ n ++  
    { a |-> m(a) + n(a) | a in set dom n inter dom m };  
  
-- 1  
mseExMul[@a]:map @a to nat*nat->map @a to nat  
mseExMul(m,n)==if n=1 then m  
    else mseCup[@a] (m,mseExMul[@a] (m,n-1));  
  
-- versao alternativa  
mseExMul2[@a]:map @a to nat*nat->map @a to nat  
mseExMul2(m,n)=={a|->m(a)*n|a in set dom m};  
  
-- 2  
mseDiff[@a]:map @a to nat*map @a to nat->map @a to nat  
mseDiff(m1,m2)=={a|->(m1(a)-m2(a))|a in set dom m1  
    ,a in set dom m2  
    &m1(a)-m2(a)>0}  
    munion {a|->m1(a)|a in set dom m1  
    &not(a in set dom m2)};  
  
-- 3  
mseCUP[@a]:seq of map @a to nat->map @a to nat  
mseCUP(s)==if len s=1 then hd s  
    else mseCup[@a] (hd s,mseCUP[@a] (tl s))  
pre len s>=1;  
  
values  
m1={"a"|->3,"b"|->3,"c"|->1,"d"|->1};  
m2={"a"|->3,"c"|->4,"e"|->1,"g"|->5};
```

5. Problem 8 - Hash tables “are” Mappings

Hash tables are well known data structures whose purpose is to efficiently combine the advantages of both static and dynamic storage of data. Static structures such as *arrays* provide random access to data but have the disadvantage of filling too much primary storage. Dynamic, pointer-based structures (*e.g.* search lists, search trees, *etc.*) are more versatile with respect to storage requirements but access to data is not as immediate.

The idea of *hashing* is suggested by the informal meaning of the term itself: a large database file is “hashed” into as many “pieces” as possible, each of which is randomly accessed. Since each sub-database is smaller than the original, the time spent on accessing data is shortened by some order of magnitude. Random access is normally achieved by a so-called *hash function*,

```
hash : Data -> Location
```

which computes, for each data item, its location in the *hash table*. Standard terminology regards as *synonyms* all data competing for the same location. A set of synonyms is called a *bucket*. There are several ways in which data collision is handled, *e.g.* *linear probing* or *overflow handling*. Below we consider the latter.

1. Write and validate a VDM-SL model *HTable* for hash-tables (with overflow handling) in which you should model (a) locations as integers; (b) collision buckets as finite sets of (data) synonyms; (c) hash tables as mappings from locations to collision buckets.
2. Assuming some predefined hash function *hash*, add to *HTable* an invariant stating that all data in the same collision bucket share the same hash-index. Can you infer from this invariant that buckets are mutually disjoint?
3. Specify the following functionality on top of *HTable* so that your invariant (above) is preserved:

```
Init : () -> HTable  
  
Insert : Data * HTable -> HTable  
  
Find   : Data * HTable -> bool  
  
Remove : Data * HTable -> HTable
```

4. Specify a function **rep** able to represent any set of **Data** as a **HTable**.
5. Further specify a left inverse **abs** of **rep**, *i.e.* a function able to retrieve the collection of data which is stored into a hash-table.

5.1. Pretty-print

types

53.0 *Data* = token;

54.0 *Locations* = \mathbb{Z} ;

55.0 *Buckets* = *Data*-set;

56.0 $HTable = Locations \xrightarrow{m} Buckets$

.1 $\text{inv } h \triangleq \forall n \in \text{dom } h \cdot \forall a \in h(n) \cdot \text{hash}(a) = n$

functions

57.0 $\text{hash} : Data \rightarrow \mathbb{Z}$

.1 $\text{hash}(x) \triangleq$

.2 is not yet specified;

58.0 $Init : () \rightarrow HTable$

.1 $Init() \triangleq$

.2 $\{\mapsto\};$

59.0 $Insert : Data \times HTable \rightarrow HTable$

.1 $Insert(x, h) \triangleq$

.2 let $aux = \text{if } \text{hash}(x) \in \text{dom } h$

.3 then $h(\text{hash}(x))$

.4 else $\{\}$ in

.5 $h \uparrow \{\text{hash}(x) \mapsto \{x\} \cup aux\};$

60.0 $Find : Data \times HTable \rightarrow \mathbb{B}$

.1 $Find(x, h) \triangleq$

.2 let $aux = h(\text{hash}(x))$ in

.3 $x \in aux;$

61.0 $Remove : Data \times HTable \rightarrow HTable$

.1 $Remove(x, h) \triangleq$

.2 let $aux = h(\text{hash}(x))$ in

.3 $h \uparrow \{\text{hash}(x) \mapsto aux \setminus \{x\}\}$

.4 pre $Find(x, h);$

62.0 $rep : Data\text{-set} \rightarrow HTable$

.1 $rep(s) \triangleq$

.2 $\{\text{hash}(x) \mapsto \{y \mid y \in s \cdot \text{hash}(y) = \text{hash}(x)\} \mid x \in s\};$

63.0 $abs1 : HTable \rightarrow Data\text{-set}$

.1 $abs1(h) \triangleq$

.2 $\bigcup \{h(a) \mid a \in \text{dom } h\}$

5.2. Source code

```
-----  
-- Rui Carlos A. Goncalves  
-- LMCC - MFP I  
-- 2006/07  
-- Problema 8  
-----  
  
types  
Data=token;  
  
--1,2  
Locations=int;  
Buckets=set of Data;  
  
HTable=map Locations to Buckets  
inv h==forall n in set dom h & forall a in set h(n) & hash(a)=n;  
  
functions  
hash:Data->int  
hash(x)==is not yet specified;  
  
--3  
Init:()->HTable  
Init()=={|->};  
  
Insert:Data*HTable->HTable  
Insert(x,h)==let aux=if hash(x) in set dom h  
                  then h(hash(x))  
                  else {}  
                  in h++{hash(x)|->{x} union aux};  
  
Find:Data*HTable->bool  
Find(x,h)==let aux=h(hash(x))  
            in x in set aux;  
  
Remove:Data*HTable->HTable  
Remove(x,h)==let aux=h(hash(x))  
              in h++{hash(x)|->aux\{x}}  
pre Find(x,h);  
  
--4  
rep:set of Data->HTable  
rep(s)=={hash(x)|->{y|y in set s & hash(y)=hash(x)}|x in set s};  
  
--5  
abs1:HTable->set of Data  
abs1(h)==dunion {h(a)|a in set dom h};
```

6. Problem 10 - A Naive Model of the WWW

Consider the following VDM-SL model specifying, at abstract level, the structure of an information system based on the 'World Wide Web' over the INTERNET, where `Ref` (page address) is a datatype of which no further details are required:

```
WWW = map Ref to URL;           -- (URL=Universal Resource Location)
URL  = seq of Unit;
Unit = PlainText | HyperLink;
PlainText = seq of Word;
Word = seq of char;
HyperLink :: link: Ref
            txt: PlainText;    -- "underlined text"
```

1. Add an invariant to `WWW` ensuring that no URL mentions a non-existing URL (NB: although this cannot be ensured for the WWW as a whole, it makes sense for the particular fragment which embodies our information system).
2. Assuming

```
wc : seq of char -> nat
-- counts the number of words in a string
```

specify

```
nrofwd : URL -> nat
-- counts the number of words in a URL
```

3. Specify

```
refsTo : Ref * WWW -> set of Ref
-- retrieves the addresses of all URLs which point at Ref
```

4. Search engines on the WWW are based on text inversion, that is, on a function

```
invert : WWW -> map Word to map Ref to nat1
```

which computes, for every word, the URLs which mention it, weighted by the number of occurrences, ie. a multiset. Specify `invert`. (NB: this inversion operation is far more elaborate in practice!)

6.1. Pretty-print

types

64.0 $Ref = \mathbb{Z}$;

65.0 $HyperLink :: link : Ref$
.1 $txt : PlainText$;

66.0 $Word = \text{char}^*$;
67.0 $PlainText = Word^*$;
68.0 $Unit = PlainText \mid HyperLink$;
69.0 $URL = Unit^*$;
70.0 $WWW = Ref \xrightarrow{m} URL$
.1 $\text{inv } www \triangleq \forall x \in \text{refs}(\text{rng } www) \cdot x \in \text{dom } www$

functions

71.0 $\text{refsaux} : Unit \rightarrow \text{Ref-set}$
.1 $\text{refsaux}(u) \triangleq$
.2 if $\text{is-HyperLink}(u)$
.3 then $\{u.\text{link}\}$
.4 else $\{\}$;

72.0 $\text{refs} : URL\text{-set} \rightarrow \text{Ref-set}$
.1 $\text{refs}(urls) \triangleq$
.2 let $x1 = \{\text{elems}(x) \mid x \in urls\}$,
.3 $x2 = \bigcup x1$,
.4 $x3 = \{\text{refsaux}(x) \mid x \in x2\}$ in
.5 $\bigcup x3$;

73.0 $wc : \text{char}^* \rightarrow \mathbb{N}$
.1 $wc(x) \triangleq$
.2 is not yet specified;

74.0 $wcunit : Unit \rightarrow \mathbb{N}$
.1 $wcunit(unit) \triangleq$
.2 if $\text{is-HyperLink}(unit)$
.3 then $\text{sum}([\text{wc}(unit.\text{txt}(i)) \mid i \in \text{inds } unit.\text{txt}])$
.4 else $\text{sum}([\text{wc}(unit(i)) \mid i \in \text{inds } unit])$;

75.0 $\text{sum} : \mathbb{N}^* \rightarrow \mathbb{N}$
.1 $\text{sum}(s) \triangleq$
.2 if $\text{len } s = 0$
.3 then 0
.4 else $\text{hd}(s) + \text{sum}(\text{tl}(s))$;

76.0 $nrofwid : URL \rightarrow \mathbb{N}$

- .1 $nrofwid(url) \triangleq$
- .2 $sum([wcunit(url(i)) \mid i \in inds\ url]);$

77.0 $refsTo : Ref \times WWW \rightarrow Ref\text{-set}$

- .1 $refsTo(ref, www) \triangleq$
- .2 $\{r \mid r \in \text{dom } www \cdot ref \in refs(\{www(r)\})\}$
- .3 **pre** $ref \in \text{dom } www$;

78.0 $extract : Unit \rightarrow Word^*$

- .1 $extract(x) \triangleq$
- .2 **if** $is\ HyperLink(x)$
- .3 **then** $x.txt$
- .4 **else** x ;

79.0 $join : URL \rightarrow Word^*$

- .1 $join(url) \triangleq$
- .2 $conc[extract(url(i)) \mid i \in inds\ url];$

80.0 $words : WWW \rightarrow Ref \xrightarrow{m} Word^*$

- .1 $words(x) \triangleq$
- .2 $\{a \mapsto join(x(a)) \mid a \in \text{dom } x\};$

81.0 $count : Word \times Ref \xrightarrow{m} Word^* \rightarrow Ref \xrightarrow{m} \mathbb{N}_1$

- .1 $count(w, m) \triangleq$
- .2 **let** $r = \{x \mapsto \text{len} [(m(x))(i) \mid i \in inds\ m(x) \cdot (m(x))(i) = w] \mid$
- .3 $x \in \text{dom } m\}$ **in**
- .4 $r \triangleright \{x \mid x \in \text{rng } r \cdot x > 0\};$

82.0 $invert : WWW \rightarrow Word \xrightarrow{m} Ref \xrightarrow{m} \mathbb{N}_1$

- .1 $invert(www) \triangleq$
- .2 **let** $ws = words(www),$
- .3 $wd = \bigcup \{elems(a) \mid a \in \text{rng } ws\}$ **in**
- .4 $\{x \mapsto count(x, ws) \mid x \in wd\}$

values

```

83.0  www : WWW = {1 ↦ [{"abc", "cde"}, {"cde", "efg"}],
      .1      2 ↦ [{"hij", "abc", "xyz"}, {"xpto"}, {"abc", "cde"}],
      .2      3 ↦ [{"hij", "xpto", "www"}, {"www", "xpto", "net"}],
      .3      4 ↦ [{"a", "b"}],
      .4      5 ↦ [{"a", "b"}],
      .5      6 ↦ [mk-HyperLink (2, [{"aaaa", "bbbb"}]), mk-HyperLink (3, [{"cccc"}])],
      .6      7 ↦ [mk-HyperLink (3, [{"aaaa"}])]

```

6.2. Source code

```

-----
-- Rui Carlos A. Goncalves
-- LMCC - MFP I
-- 2006/07
-- Problema 10
-----

types
Ref=int;

HyperLink::link:Ref
      txt :PlainText;

Word=seq of char;

PlainText=seq of Word;

Unit=PlainText|HyperLink;

URL=seq of Unit;

WWW=map Ref to URL
inv www==forall x in set refs(rng www) & x in set dom www;

functions
--1
refsaux:Unit->set of Ref
refsaux(u)==if is_HyperLink(u)
      then {u.link}
      else {};

refs:set of URL->set of Ref
refs(urls)==let x1={elems(x)|x in set urls},
      x2=dunion x1,
      x3={refsaux(x)|x in set x2}
      in dunion x3;

--2
wc:seq of char->nat
wc(x)==is not yet specified;

wcunit:Unit->nat

```



```

wcunit(unit)==if is_HyperLink(unit)
    then sum([wc(unit.txt(i))|i in set inds unit.txt])
    else sum([wc(unit(i))|i in set inds unit]);

sum:seq of nat->nat
sum(s)==if len s=0
    then 0
    else hd(s)+sum(tl(s));

nrofw:URL->nat
nrofw(url)==sum([wcunit(url(i))|i in set inds url]);

--3
refsTo:Ref*WWW->set of Ref
refsTo(ref,www)={r|r in set dom www & ref in set refs({www(r)})}
pre ref in set dom www;

--4
extract:Unit->seq of Word
extract(x)==if is_HyperLink(x)
    then x.txt
    else x;

join:URL->seq of Word
join(url)==conc [extract(url(i))|i in set inds url];

words:WWW->map Ref to seq of Word
words(x)={a|->join(x(a))|a in set dom x};

count:Word*map Ref to seq of Word->map Ref to nat1
count(w,m)==let r={x|->len [(m(x))(i)|i in set inds m(x) & (m(x))(i)=w]
    |x in set dom m}
    in r:>{x|x in set rng r & x>0};

invert:WWW->map Word to map Ref to nat1
invert(www)==let ws=words(www),
    wd=dunion {elems(a)|a in set rng ws}
    in {x|->count(x,ws)|x in set wd};

values
www:WWW={1|->[["abc","cde"],["cde","efg"]],
    2|->[["hij","abc","xyz"],["xpto"],["abc","cde"]],
    3|->[["hij","xpto","www"],["www","xpto","net"]],
    4|->[["a","b"]],
    5|->[["a","b"]],
    6|->[mk_HyperLink(2,["aaaa","bbbb"]),mk_HyperLink(3,["cccc"])],
    7|->[mk_HyperLink(3,["aaaa"])]
};

```

7. Problem 11 - Modelling a Memory Cache

Analyse the formal VDM-SL model below of a memory *cache*. *Try to improve it* and design a test suite in the toolbox to check its behaviour.

```
types
  Addr = token;
  Data = token;

  Memory :: addr: set of Addr
          mapx: map Addr to Data;

  MainMemory = Memory;

  Cache :: addr: set of Addr
         mapx: map Addr to Data
         dirty: set of Addr;

  System :: cache: Cache
          main: MainMemory
          inv system == system.cache.addr subset system.main.addr;

functions

  badAddr: System -> set of Addr
  badAddr (s) == { x | x in set s.cache.addr &
                  s.cache.mapx(x) <> s.main.mapx(x) };

  load: System * Addr -> System
  load (s, a) == mk_System(
    mk_Cache(s.cache.addr union {a},
             s.cache.mapx munion {a|->s.main.mapx(a)},
             s.cache.dirty),
    s.main)
  pre a not in set s.cache.addr;

  systemWrite: System * Data * Addr -> System
  systemWrite (s, d, a) ==
    mk_System(
      mk_Cache(s.cache.addr,
               s.cache.mapx ++ {a|->d},
               s.cache.dirty union {a}),
      s.main);

  flush: System -> System
  flush (s) == let x in set s.cache.addr
    in mk_System(
      mk_Cache(s.cache.addr \ {x},
               {x} <-: s.cache.mapx,
               s.cache.dirty \ {x}),
      mk_Memory(s.main.addr,
                s.main.mapx ++ {x|->s.cache.mapx(x)}));
```

7.1. Pretty-print

types

```

84.0  Addr = token;

85.0  Data = token;

86.0  Memory :: addr : Addr-set
      .1      mapx : Addr  $\xrightarrow{m}$  Data
      .2      inv m  $\triangleq$  dom m.mapx  $\subseteq$  m.addr;

87.0  MainMemory = Memory;

88.0  Cache :: addr : Addr-set
      .1      mapx : Addr  $\xrightarrow{m}$  Data
      .2      dirty : Addr-set
      .3      inv c  $\triangleq$ 
c.dirty  $\subseteq$  c.addr  $\wedge$ 
dom c.mapx  $\subseteq$  c.addr;

89.0  System :: cache : Cache
      .1      main : MainMemory
      .2      inv system  $\triangleq$  system.cache.addr  $\subseteq$  system.main.addr

```

functions

```

90.0  badAddrs : System  $\rightarrow$  Addr-set
      .1  badAddrs (s)  $\triangleq$ 
      .2  {x | x  $\in$  s.cache.addr  $\cdot$ 
      .3  s.cache.mapx (x)  $\neq$  s.main.mapx (x)};

91.0  load : System  $\times$  Addr  $\rightarrow$  System
      .1  load (s, a)  $\triangleq$ 
      .2  mk-System
      .3  (
      .4  mk-Cache (s.cache.addr  $\cup$  {a},
      .5  s.cache.mapx  $\sqcup$  {a  $\mapsto$  s.main.mapx (a)},
      .6  s.cache.dirty),
      .7  s.main)
      .8  pre a  $\notin$  s.cache.addr ;

```

92.0 $systemWrite : System \times Data \times Addr \rightarrow System$

```
.1  $systemWrite (s, d, a) \triangleq$   
.2    $mk\text{-}System$   
.3     (  
.4        $mk\text{-}Cache (s.cache.addr, s.cache.mapx \uparrow \{a \mapsto d\},$   
.5          $s.cache.dirty \cup \{a\}),$   
.6        $s.main)$   
.7  
.8  $pre\ a \in s.cache.addr ;$ 
```

93.0 $flush : System \rightarrow System$

```
.1  $flush (s) \triangleq$   
.2    $let\ x \in s.cache.addr\ in$   
.3      $mk\text{-}System$   
.4       (  
.5          $mk\text{-}Cache (s.cache.addr \setminus \{x\},$   
.6            $\{x\} \Leftarrow s.cache.mapx,$   
.7            $s.cache.dirty \setminus \{x\}),$   
.8          $mk\text{-}Memory (s.main.addr,$   
.9            $s.main.mapx \uparrow \{x \mapsto s.cache.mapx(x)\})$   
.10   $pre\ s.cache.addr \neq \{\}$ 
```

values

94.0 $a0 : Addr = mk\text{-}token (0);$

95.0 $a1 : Addr = mk\text{-}token (1);$

96.0 $a2 : Addr = mk\text{-}token (2);$

97.0 $a3 : Addr = mk\text{-}token (3);$

98.0 $a4 : Addr = mk\text{-}token (4);$

99.0 $a5 : Addr = mk\text{-}token (5);$

100.0 $a6 : Addr = mk\text{-}token (6);$

```

101.0  a7 : Addr = mk-token (7);

102.0  a8 : Addr = mk-token (8);

103.0  a9 : Addr = mk-token (9);

104.0  d0 : Data = mk-token ("");

105.0  d1 : Data = mk-token ("a");

106.0  d2 : Data = mk-token ("b");

107.0  d3 : Data = mk-token ("c");

108.0  d4 : Data = mk-token ("d");

109.0  d5 : Data = mk-token ("e");

110.0  d6 : Data = mk-token ("f");

111.0  d7 : Data = mk-token ("g");

112.0  d8 : Data = mk-token ("h");

113.0  d9 : Data = mk-token ("i");

114.0  sys : System = mk-System
      .1      (
      .2          mk-Cache ({} ,
      .3              {↦} ,
      .4              {}),
      .5          mk-Memory ({a0, a1, a2, a3, a4, a5, a6, a7, a8, a9},
      .6              {a0 ↦ d0, a1 ↦ d0, a2 ↦ d0, a3 ↦ d0, a4 ↦ d0,
      .7              a5 ↦ d0, a6 ↦ d0, a7 ↦ d0, a8 ↦ d0, a9 ↦
d0}));

```

```
115.0  sys1 = systemWrite (load (sys, a3), d1, a3);

116.0  sys2 = systemWrite (load (sys1, a2), d8, a2);

117.0  sys3 = systemWrite (load (sys2, a7), d4, a7);

118.0  sys4 = systemWrite (load (sys3, a4), d2, a4);

119.0  sys5 = flush (sys4);

120.0  sys6 = flush (sys5);

121.0  sys7 = flush (sys6)
```

7.2. Source code

```
-----
-- Rui Carlos A. Goncalves
-- LMCC - MFP I
-- 2006/07
-- Problema 11
-----

types
Addr = token;
Data = token;

Memory :: addr: set of Addr
        mapx: map Addr to Data
inv m == dom m.mapx subset m.addr;

MainMemory = Memory;

Cache :: addr: set of Addr
        mapx: map Addr to Data
        dirty: set of Addr
inv c == c.dirty subset c.addr
        and dom c.mapx subset c.addr;

System :: cache: Cache
        main: MainMemory
inv system == system.cache.addr subset system.main.addr;
```

```

functions

badAddr: System -> set of Addr
badAddr (s) == { x | x in set s.cache.addr &
                s.cache.mapx(x) <> s.main.mapx(x) };

load: System * Addr -> System
load (s, a) == mk_System(
    mk_Cache(s.cache.addr union {a},
            s.cache.mapx munion {a|->s.main.mapx(a)},
            s.cache.dirty),
    s.main)
pre a not in set s.cache.addr;

systemWrite: System * Data * Addr -> System
systemWrite (s, d, a) ==mk_System(
    mk_Cache(s.cache.addr,
            s.cache.mapx ++ {a|->d},
            s.cache.dirty union {a}),
    s.main)
pre a in set s.cache.addr;

flush: System -> System
flush (s) == let x in set s.cache.addr
    in mk_System(
        mk_Cache(s.cache.addr \ {x},
                {x} <-: s.cache.mapx,
                s.cache.dirty \ {x}),
        mk_Memory(s.main.addr,
                s.main.mapx ++ {x|->s.cache.mapx(x)}))
pre s.cache.addr<>{};

values
a0:Addr=mk_token(0);
a1:Addr=mk_token(1);
a2:Addr=mk_token(2);
a3:Addr=mk_token(3);
a4:Addr=mk_token(4);
a5:Addr=mk_token(5);
a6:Addr=mk_token(6);
a7:Addr=mk_token(7);
a8:Addr=mk_token(8);
a9:Addr=mk_token(9);

d0:Data=mk_token("");
d1:Data=mk_token("a");
d2:Data=mk_token("b");
d3:Data=mk_token("c");
d4:Data=mk_token("d");
d5:Data=mk_token("e");
d6:Data=mk_token("f");
d7:Data=mk_token("g");
d8:Data=mk_token("h");

```

```
d9:Data=mk_token("i");

sys:System=mk_System(
    mk_Cache({},
        {|->},
        {}
    ),
    mk_Memory({a0,a1,a2,a3,a4,a5,a6,a7,a8,a9},
        {a0|->d0,a1|->d0,a2|->d0,a3|->d0,a4|->d0,
        a5|->d0,a6|->d0,a7|->d0,a8|->d0,a9|->d0
        }
    )
);

sys1=systemWrite(load(sys,a3),d1,a3);
sys2=systemWrite(load(sys1,a2),d8,a2);
sys3=systemWrite(load(sys2,a7),d4,a7);
sys4=systemWrite(load(sys3,a4),d2,a4);
sys5=flush(sys4);
sys6=flush(sys5);
sys7=flush(sys6);
```