Phase 4. Code Generation The Compildres

Harold Treen
Paul Wells

April 6th, 2015

1 - Updating Tokens

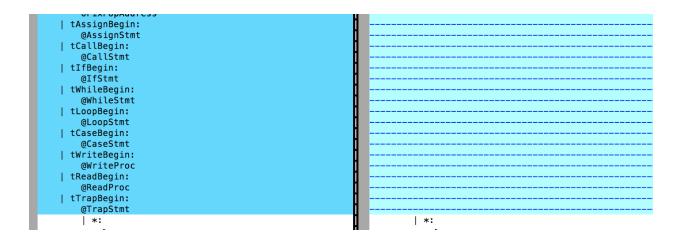
As with all previous phases, the input and output tokens had to be updated in coder.ssl. Tokens were added for the new string operations and loop syntax. Old tokens for char operations were removed.

```
% Added non-compound T-Codes
          tFetchString
   86
          tAssignString
   87
          tStoreParmString
   88
          tSubscriptString
   89
          tConcatenate
   90
   91
          tSubstring
   92
93
94
          tLength
tStringEqual
          tLoopBegin
   95
          tLoopBreakIf
   96
          tCaseDefault
   97
          tCaseElse
          % Compound T-codes are those that take operands
   98
          tLiteralAddress
   99
          firstCompoundToken = tLiteralAddress
  100
         tLineNumber
117
          % Added Compound T-Codes
         tLoopTest
118
         tLoopEnd
119
120
121
         % Return the top entry's jump condition (kind field)
288
289
290
         oOperandPushString
         % Push an entry with addressing mode manifest, length w
         % a value set to the data area address of the string re
        tFetchInteger
 56
        tFetchChar
 57
        tFetchBoolean
 58
        tAssignBegin
 59
        tAssignAddress
 60
        tAssignInteger
 62
        tAssignChar
 63
        tAssignBoolean
 64
        tStoreParmAddress
 65
        tStoreParmInteger
 66
        tStoreParmChar
 67
        tStoreParmBoolean
        tSubscriptBegin
 68
        tSubscriptAddress
 69
        tSubscriptInteger
tSubscriptChar
 70
        tSubscriptBoolean
 72
 73
        tArrayDescriptor
        tFileDescriptor
 74
 75
        tIfBegin
        tIfEnd
 76
 77
        tCaseBegin
 78
        tWhileBegin
 79
        tRepeatBegin
 80
        tRepeatControl
 81
        tCallBegin
 82
        tParmEnd
        tProcedureEnd
 83
        tWriteBegin
 84
        tReadBegin
 85
        tTrapBegin
 86
        tWriteEnd
 87
        tReadEnd
 88
 89
        % Compound T-codes are those that take operands
 90
        tLiteralAddress
 91
        firstCompoundToken = tLiteralAddress
 92
        tLiteralInteger
        tLiteralChar
        tLiteralBoolean
```

Once the tokens in the .ssl file were updated, a .def file was generated which was used to update definitions in coder.pt.

2 - Drift Modules

Since modules are essentially a scoping construct, most of the work for modules has already been done with one exception: Now, Statements do not strictly have to follow Declarations. Therefore in the Block rule we added the choice alternatives for statements into the loop with declarations.



3 - Drift Loop Templates

For Drift loops the first step was to remove the old RepeatStmt Rule. The RepeatStmt rule was replaced with a LoopStmt rule. As you can see below, the diffs for this piece of code are messy.



You can see that the RepeatStmt was removed but to show that the LoopStmt rule has been added we'll just show code snippets of our current code instead of a diff for clarity.

```
LoopStmt:
   % Save the target address for the top-of-loop branch
   oFixPushAddress
                              % bNNN:
   oEmitMergeSourceCoordinate
   % Added Statements before break condition
   @Statements
   % Require break if
   tLoopBreakIF
@OperandPushBooleanControlExpression % .... cond
% Changed from tWhileTest to tLoopTest
ntstLoopTest
on:%oOperandComplementJumpCondition
                                       % ... !cond
es % Optimize if the condition is know at compile time
   [ oOperandChooseJumpCondition
ze LiJnever:
indChooseJ% Exit condition is always false (while true) - an infinite loop.
           % emit no conditional branch, just fall into loop body
% Exit coroFixAndFreeFalseBranchesse (until false)
oFik/iJalways:alseBranches
o Emit U% c Exiti condition cis always true (while false) - a nop.
       % flush the loop body
lalwaysoFixAndFreeFalseBranches
       @SkipToEndWhile always true (until
      oFixPopAddress out backward branch
                                 % fNN:, just in case
oFixAndFreeShuntLists
       oOperandPop
       >>
       % Emit a conditional forward branch to exit the loop.
       % True (inverted false) branches follow the conditional
       % exit path, false branches fall through to the loop body.
       @OperandInfixOr ----% - j!cond fNNN
   Э.
   @Statements
                              % loop body
   % Changed from tWhileEnd to tLoopEnd
   tLoopEnd
 d oEmitUnconditionalBackwardBranch
                                              jmp bNNN
```

oFixPopAddress oFixAndFreeShuntList oOperandPop;

% fNNN:

You can see that the code for a loop statement is similar to the code for a while loop except that statements are added before the break condition and of course tWhileEnd is changed to tLoopEnd.

Also, the choice for tRepeatBegin had to be changed to tLoopBegin in the Block and Statement rules like so:



Finally, in coder.pt, tRepeatTest had to be replaced with tLoopTest and tLoopEnd inside the code which accepts compound tokens.



4 - Drift Switch Template

The only change that was required for drift switch statements was to edit the EmitDefaultCaseAbort rule so that if there is a tCaseElse token (indicating that there is a default case) then to execute that rule when no case matches are found instead of aborting the case statement.



5 - String Templates

5.1 - Literal Strings

tLiteralString allows string variables to be declared within the machine code. It replaced the tStringDescriptor cases from the old code generator. The choice was added to both the regualr "OperandPushExpression" and optimized "OperandPushExpressionAssignPopPop" rule.

```
% Added choice for tLiteralString
1506
           tLiteralString:
1507
              % Logic is the same except no tStringDescriptor acc
1508
              % Emit string literal to data area
1509
             oEmitNone(iData)
                                                .data
1510
              % Emit the string
1511
             oEmitString
                                       % sNNN: .asciz "SSSSS"
1512
             oEmitNone(iText)
1513
                                                .text
              % Get the string literal's address
1514
             oOperandPushString
1515
              @EmitStringDescriptor
                                                lea sNNN, %T
1516
```

coder.ssl

5.2 - Fetching String

Fetching strings was done by forcing the address of the string operand into a temp variable using the "EmitStringDescriptor" rule. It replaced the tFetchChar operation.

coder.ssl

5.3 - Assign Strings

String assignment was added as a case in "OperandPushExpressionAssignPopPop". The new rule saves all temporary registers, puts the string addresses in registers and does a trap call to move the contents at the first address to the address of the second. Both operands are then popped from the stack.

```
2182
          % Add tAssignString
 2183
           tAssignString:
              @OperandAssignStringPopPop
 2184
2185
3132 OperandAssignStringPopPop:
         % Generate code to assign the right (top) operand's val
3133
         % the left (second) operand and pop both operands. For
3134
3135
         @SaveTempRegsToStack
3136
         @EmitStringDescriptor %lea s2, %T
         @OperandForceToStack %pushl %T
3137
         @OperandPopAndFreeTemp
3138
         @EmitStringDescriptor %lea s1,
3139
         @OperandForceToStack %pushl, %T
3140
3141
         @OperandPopAndFreeTemp
3142
         % Make trap call
3143
         oOperandPushMode(mTrap)
3144
3145
         oOperandSetValue(trAssignString)
3146
         oOperandSetLength(word)
         oEmitSingle(iCall)
3147
3148
         oOperandPop
3149
3150
         % Pop arguments
         oOperandPushMode(mStackReg)
3151
3152
         oOperandSetLength(word)
         oOperandPushMode(mManifest)
3153
         oOperandSetLength(word)
3154
3155
         oOperandSetValue(eight)
3156
         oEmitDouble(iAdd)
3157
         oOperandPop
         oOperandPop
3158
3159
         @RestoreTempRegsFromStack;
3160
3161
```

coder.ssl

5.4 - Parameter Strings

String parameter handling was added so that strings could be passed to functions. The added code moves the string value into a variable which is accessible by the function.

coder.ssl

5.5 - String Subscripting

String subscripting was done identically to integer subscripting. The difference was a small change so that the calculated offsets would be multiplied by 256 (the size of strings) by multiplying in coder.pt and shifting in coder.ssl.

```
1701 OperandSubscriptNonManifestStringPop:
1702 [ oOptionTestChecking
1703 | yes:
1704 @OperandCheckedSubscriptNonManifestStringPop
1705 | *:
1706 @OperandUncheckedSubscriptNonManifestStringPop
1707 ];
1708
```

```
OperandUncheckedSubscriptNonManifestStringPop:
 1968
            % Optimized non-bounds checking subscript operation
                                                   % ... subscript, arraydesc
            oOperandSwap
 1969
1970
               oOperandChooseMode
 1971
               mTempIndirect:
 1972
                 % Var parameter array - don't know the characteristics
% until run time, so give up and use regular checked subscripting
1973
1974
1975
                 oOperandSwap
                  @OperandCheckedSubscriptNonManifestStringPop
 1976
1977
            mStatic:
1978
1979
                  % Any other array - know all the characteristics now,
1980
                  % so optimize subscripting as best we can
 1981
                 oOperandSwap
                                                   % ... arraydesc, subscript
1982
                 % Scale subscript by string element size 
@OperandForceIntoTemp % mov1 ;
oOperandPushMode(mManifest)
1983
                                                                         subscript, %T
 1984
 1985
                 oOperandSetLength(word)
1986
                 oOperandSetValue(eight)
1987
                 oEmitDouble(iShl)
                                                              shl
                                                                         $2, %T
1988
 1989
                 oOperandPop
                                                   % ... %T, arraydesc
1990
                 oOperandSwap
1991
                  % Fold lower bound into array address to avoid normalizing
1992
                  % subscript at run time
                 1994
1995
1996
 1997
 1998
                 oOperandSetLength(word)
oOperandSetValue(eight)
1999
                                                  es % ... %T, lower, arraydesc+8, 8
% ... %T, lower, arraydesc+8
% ... %T, arraydesc+8, lower
% ... %T, arraydesc+8, lower, lower
es % (scale lower bound by integer size
                 oOperandAddManifestValues
2000
                  oOperandPop
2001
2002
                 oOperandSwap
                 oOperandPushCopy
oOperandAddManifestValues
2003
2004
                 oOperandPop
2005
                 oOperandPushCopy
oOperandAddManifestValues
2006
                 OOperandAddManifestValues % ... %T, arraydesc+8, lower*4, lower*4 oOperandPop % ... %T, arraydesc+8, lower*4 oOperandSubtractManifestValues % ... %T, arraydesc+8-lower*4, lower*4 oOperandPop % ... %T, arraydesc+8-lower*4
2007
2008
2009
2010
                 % Add array base to subscript
oOperandSetMode(mStaticManifest) % (u+normalizedArrayBase)
% addl $u+normalizedArrayBase
2011
2012
2013
                                                                        $u+normalizedArrayBase, %T
2014
2015
                 oOperandPop
2016
2017
                  % Element address is in %T
                 oOperandSetMode(mTempIndirect)
2018
                                                                   % ... (%T)
2019
                 oOperandSetLength(string)
2020
            1;
2021
```

coder.ssl

```
{ Longword (4 Dytes) on SUN }
                             wordsize = 4;
                  506
                             stringSize = 256;
                  507
                  508
                  if operandStkLength[operandStkTop-1] = word then subscript := subscript * wordSize { Convert a byte offset to a word offset }
1805
1806
1807
                  else
1809
                  begin
                       if operandStkLength[operandStkTop-1] = string then subscript := subscript * stringSize
1810
                            { Convert a byte offset to a string offset}
1811
1812
1814
```

coder.pt

5.6 - Concatenation Operation

A case for the concatenation operation was added to both the "OperandPushExpression" and "OperandPushExpressionAssignPopPop" rules. The new operation pushes two string addresses into temp registers which can then be accessed by the "trConcatenate" trap call. This operation returns the result in a temp register which is then saved and restored.

```
tConcatenate:
           @OperandConcatenatePop
     OperandConcatenatePop:
2355
          @SaveTempRegsToStack
2356
          @EmitStringDescriptor %lea s1,
2357
2358
          @OperandForceToStack
          @OperandPopAndFreeTemp
2359
          @EmitStringDescriptor %lea s1,
2360
          @OperandForceToStack %pushl, %T
2361
          @OperandPopAndFreeTemp
2362
2363
          % Make trap call
2364
          oOperandPushMode(mTrap)
2365
2366
          oOperandSetValue(trConcatenate)
          oOperandSetLength(word)
2367
          oEmitSingle(iCall) % call pttrap103
2368
         oOperandPop
2369
2370
          % Pop arguments
2371
          oOperandPushMode(mStackReg)
2372
          oOperandSetLength(word)
2373
         oOperandPushMode(mManifest)
2374
          oOperandSetLength(word)
2375
          oOperandSetValue(eight)
2376
         oEmitDouble(iAdd)
2377
          oOperandPop
2378
         oOperandPop
2379
2380
          % Save result
2381
          @SaveResult
2382
2383
          @RestoreTempRegsFromStack
2384
2385
2386
          % Move result to temp
2387
          @ResultToTemp;
2388
```

coder.ssl

5.7 - Length Operation

The length operation was also added as a case to both the "OperandPushExpression" and "OperandPushExpressionAssignPopPop" rules. The length operation pushes a string operand address into a temp register and then calls a trap which returns the length. The result is then saved and restored.

```
| tLength:

    @OperandLength
```

```
2389 OperandLength:
         @SaveTempRegsToStack
2390
         @EmitStringDescriptor %lea s1, %T
2391
2392
         @OperandForceToStack %pushl %T
2393
         @OperandPopAndFreeTemp
2394
         % Make trap call
2395
         oOperandPushMode(mTrap)
2396
2397
         oOperandSetValue(trLength)
         oOperandSetLength(word)
2398
         oEmitSingle(iCall) % call pttrap105
2399
2400
         oOperandPop
2401
2402
         % Pop Arguments
2403
         oOperandPushMode(mStackReg)
2404
         oOperandSetLength(word)
         oOperandPushMode(mManifest)
2405
2406
         oOperandSetLength(word)
2407
         oOperandSetValue(four)
         oEmitDouble(iAdd)
2408
2409
         oOperandPop
         oOperandPop
2410
2411
2412
         % Save result
         @SaveResult % movl eax , esi
2413
2414
         @RestoreTempRegsFromStack
2415
2416
         % Move result to temp
2417
         @ResultToTemp;
2418
2410
```

coder.ssl

5.8 - Chr Operation

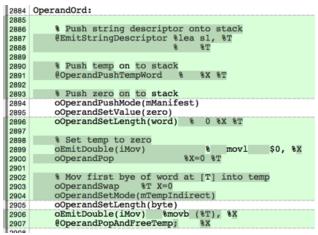
The "OperandChr" rule was updated to work with the new string type. The code now saves the integer parameter in a register before calling a trap to return the corresponding string. The result is then saved and restored.

```
2851 OperandChr:
                  @SaveTempRegsToStack
                  @OperandForceToStack % pushl il
 2854
 2855
2856
2857
2858
                  @OperandPopAndFreeTemp
                 % Make trap call
oOperandPushMode(mTrap)
oOperandSetValue(trChrString) % call pttrap102
oOperandSetLength(word)
oEmitSingle(iCall)
eOperandSetLength(word)
 2859
 2859
2860
2861
2862
2863
2864
2865
2866
                  oOperandPop
                  % Pop arguments
oOperandPushMode(mStackReg)
oOperandSetLength(word)
oOperandPushMode(mManifest)
 2867
2868
2869
2870
                  oOperandSetLength(word)
oOperandSetValue(four)
oEmitDouble(iAdd)
                                                                    % addl $4, %esp
 2871
                  oOperandPop
 2872
2873
2874
                  oOperandPop
                  % Save Result
 2875
2876
2877
2878
2879
                  @RestoreTempRegsFromStack
                  % Force result into temp
@ResultToTemp
```

coder.ssl

5.9 - Ord Operation

The ord operation needed to be updated to work with the string type. The new code moves a string address to a temp register. A trap is then called to return the corresponding integer in another temp register. The result is then saved and restored.



coder.ssl

5.10 - Equality Operation

A case for accepting the "tStringEqual" token was added to "OperandPushExpression". The case then calls "OperandStringEqualPopPop". This operation puts two string addresses into temporary registers. A trap call is then made which looks at the registers and determines if the contents at the addresses are equal. The strings are then popped from the OperandStack and the result is save then restored.

```
1552 | tStringEqual:
1553 @OperandStringEqualPop
```

```
3328 OperandStringEqualPop:
3329
3330
               % Generate code to compare string equality
               @SaveTempRegsToStack
@EmitStringDescriptor %lea s2, %T
@OperandForceToStack %pushl %T
3331
3332
3333
               @OperandPopAndFreeTemp
@EmitStringDescriptor %lea s1,
 3334
 3335
               @OperandForceToStack %push1, %T
@OperandPopAndFreeTemp
 3337
3338
3339
               % Make trap call
oOperandPushMode(mTrap)
oOperandSetValue(trStringEqual) % call pttrap106
3340
3341
               oOperandSetLength(word)
oEmitSingle(iCall)
 3342
 3343
 3344
               o0perandPop
 3345
               % Pop arguments
 3346
3347
3348
              oOperandPushMode(mStackReg)
oOperandSetLength(word)
oOperandPushMode(mManifest)
oOperandSetLength(word)
3349
3350
               oOperandSetValue(eight)
oEmitDouble(iAdd)
 3351
 3352
 3353
               oOperandPop
 3354
               oOperandPop
 3355
3356
3357
               % Save result
@SaveResult % movl eax, esi
3358
3359
               @RestoreTempRegsFromStack
3360
               % Move result into temp
@ResultToTemp;
3362
```

coder.ssl

6 - Testing

The coder phase was thoroughly tested using a variety of focused tests and provided example. An automation suite using rspec was also created to make it very simple to run tests quickly, compare outputs and find regressions.

6.1 - If tests

If statements were tested for three cases: calling the if code, calling the elsif code and calling the else code. This was done with a simple if chain where the desired block printed true and the others printed false. The automation suite checks that the generated code prints true to verify the correct behaviour.

Test Files: if.pt, elself.pt, else.pt



```
extern output
    var a: integer
    a = 2
    if a == 2
      write("true")
    else if a == 3
      write("false")
10
      write("false")
11
12
    end
13
14
    writeln
15
```

6.2 - Loop Tests

Loop statements were tested by initializing an array of strings and having a loop print out the elements and increment an index at each iteration. When the index reached the end of the array the loop exited. An expected output file was used to verify the output.

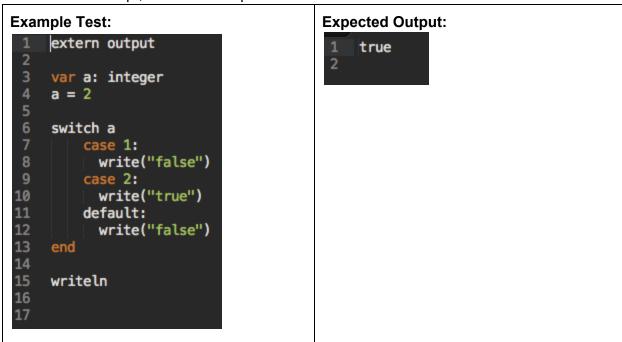
Test Files: loopTest.pt

```
Example Test:
                                                                    Expected Output:
     extern output
     var x: integer
     var s : array [1 .. 12] of string
                                                                             2
                                                                             3
                                                                             4
     s[2] = "2"
     s[3] = "3"
                                                                             5
     s[4] = "4"
                                                                             6
                                                                             7
11
12
13
14
15
16
17
18
19
20
21
22
23
24
                                                                            8
     s[8] = "8"
                                                                      9
                                                                             9
    s[10] = "10"
s[11] = "11"
                                                                     10
                                                                             10
                                                                     11
                                                                             11
                                                                     12
                                                                             12
     x=1
                                                                     13
         write(s[x])
         writeln
         x = x + 1
break if x == 13
```

6.3 - Switch Tests

Switch statements were tested for both the regular case and the default. If the correct block was executed, the word true would be printed - otherwise false.

Test Files: switch.pt, switchDefault.pt

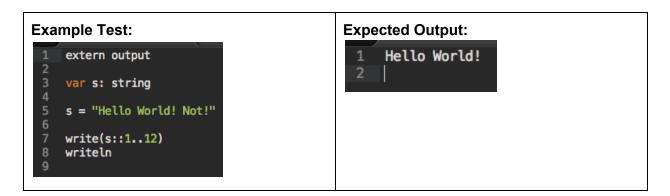


6.4 - String Tests

Strings were tested thoroughly due to all the added features. Tests covered a variety of areas including: substrings, assignment, concatenation, array subscripting, equality operation, string parameters and empty strings.

Substrings

Substrings were tested by assigning a string variable and writing only a section of it. The test would pass if the correct substring was returned.



Assignment

Assignment was tested by creating simple string variables and verifying that writing the variable produced the correct value.

```
Example Test:

1  extern output
2  var s: string
3  s = "this works"
4  write(s)
5  writeln
6  |

Expected Output:
1  |this works
2
```

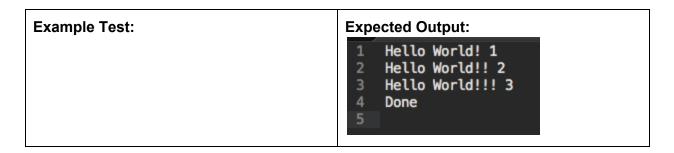
Concatenation

Concatenation was tested by adding strings together in a variety of ways (eg. literal + variable, literal + literal, etc.). For each concatenation, the result was output to be verified.

```
Example Test:
                                          Expected Output:
                                               HelloWorld
     extern output
                                               Hello World
 3 var h : string
                                               Hello World
 4 var w : string
                                               Hello World
 5 h = "Hello"
    w = "World"
 8 write(h + w)
 9 writeln
 10 write(h + " World")
 11 writeln
 12 write("Hello " + w)
 13 writeln
 14 write("Hello " + "World")
   writeln
```

Array Subscripting

Subscripting was done by creating an array of string variables and printing out the contents at each index.



```
extern output
       var sarray : array [1 .. 10] of string
      sarray[1] = "Hello "
     sarray[1] = Hetto
sarray[2] = "World"
sarray[3] = "! 1"
sarray[4] = "Hello "
     sarray[5] = "Hello"
sarray[6] = "!! 2"
sarray[7] = "Hello"
11
12
13
14
15
16
17
18
20
21
22
23
24
25
29
30
31
      sarray[8] = "World"
sarray[9] = "!!! 3"
       sarray[10] = "Done"
       write(sarray[1])
      write(sarray[2])
      write(sarray[3])
      writeln
      write(sarray[4])
write(sarray[5])
       write(sarray[6])
       writeln
      write(sarray[7])
write(sarray[8])
write(sarray[9])
       writeln
       write(sarray[10])
       writeln
```

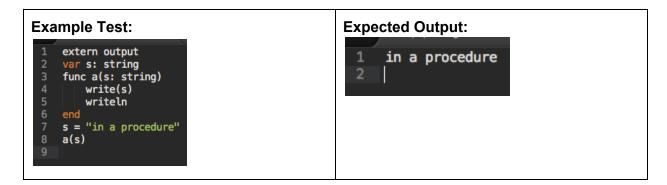
Equality Operations

String equal and not equal were tested with an if statement that would output true if the comparison output the correct result.



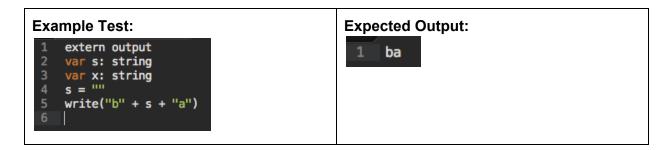
String Parameters

String parameters were tested by creating a function that would accept a string and print it. If the parameter was passed correctly, an output would show.



Empty String

The empty string was tested by adding it to other strings and ensuring the correct combined string was output.



6.5 - Chr and Ord

Chr was tested by outputting the sum of two chr operations and ensuring the correct string was produced.

Ord was tested by taking two Chrs and ensuring the correct codes were printed out.

Test Files: writeChar.pt, stringOrd.pt



6.5 - Example Tests

The example programs that only output were also added to the automation suite. These include: bubble.pt, lunch.pt, pascal.pt, primes.pt and stars.pt.

To ensure that examples with user input were not excluded, a special test case was also added for bust.pt.