PrettyRoo.hs

```
1  ------------------------------------------------------------
2  -- COMP90045 Programming Language Implementation Project --
3  --                    Roo Compiler                        --
4  --  Implemented by Xulin Yang                             --
5  --  read from the bottom to top                           --
6  ------------------------------------------------------------
7  module PrettyRoo (pp)
8  where
9  import RooAST
10 import Data.List
11
12 --  Pretty print a whole program:
13 pp :: Program -> String
14 pp = strProgram
15
16 ------------------------------------------------------------------
17 --  helper functions
18 ------------------------------------------------------------------
19
20 -- add space indentations based on the input indentation level
21 addIndentation :: Int -> String
22 addIndentation 0 = ""
23 addIndentation n = "    " ++ addIndentation (n - 1)
24
25 newline :: String
26 newline = "\n"
27
28 semicolon :: String
29 semicolon = ";"
30
31 comma :: String
32 comma = ", "  -- There should be a single space after a comma
33
34 surroundByParens :: String -> String
35 surroundByParens s = "(" ++ s ++ ")"
36
37 surroundByBrackets :: String -> String
38 surroundByBrackets s = "[" ++ s ++ "]"
39
40 ------------------------------------------------------------------
41 --  AST toString functions
42 ------------------------------------------------------------------
43
44 strBaseType :: BaseType -> String
45 strBaseType BooleanType = "boolean"
46 strBaseType IntegerType = "integer"
47
48 strDataType :: DataType -> String
49 strDataType (AliasDataType t) = t -- t is String already
50 strDataType (BasyDataType b) = strBaseType b
51
```

```haskell
strBooleanLiteral :: BooleanLiteral -> String
strBooleanLiteral True = "true"
strBooleanLiteral False = "false"

-- StringLiteral can be directly used as it is String type

-- IntegerLiteral can be turned to string by show as it is Int type

-----------------------------------------------------------------
-- An lvalue (<lvalue>) has four (and only four) possible forms:
--     <id>
--     <id>.<id>
--     <id>[<exp>]
--     <id>[<exp>].<id>
-----------------------------------------------------------------
strLValue :: LValue -> String
strLValue (LId ident) = ident
strLValue (LDot ident1 ident2) = ident1 ++ "." ++ ident2
strLValue (LBrackets ident exp) = ident ++ (surroundByBrackets (strExp exp))
strLValue (LBracketsDot ident1 exp ident2) = ident1 ++ (surroundByBrackets (strExp exp)) ++ "." ++ ident2

-----------------------------------------------------------------
-- Exp related toString & helper functions (with parens elimination)
-----------------------------------------------------------------
-- exp1 has higher precendence than exp2, higher if exp2 has no operator
isHigherPrecendence :: Exp -> Exp -> Bool
isHigherPrecendence exp1@(Op_neg _) exp2 =
  case exp2 of
    (Op_neg _  ) -> False
    _            -> True
isHigherPrecendence exp1@(Op_mul _ _) exp2 =
  case exp2 of
    (Op_neg _  ) -> False
    (Op_mul _ _) -> False
    (Op_div _ _) -> False
    _            -> True
isHigherPrecendence exp1@(Op_div _ _) exp2 =
  case exp2 of
    (Op_neg _  ) -> False
    (Op_mul _ _) -> False
    (Op_div _ _) -> False
    _            -> True
isHigherPrecendence exp1@(Op_add _ _) exp2 =
  case exp2 of
    (Op_neg _  ) -> False
    (Op_mul _ _) -> False
    (Op_div _ _) -> False
    (Op_add _ _) -> False
    (Op_sub _ _) -> False
    _            -> True
isHigherPrecendence exp1@(Op_sub _ _) exp2 =
  case exp2 of
    (Op_neg _  ) -> False
    (Op_mul _ _) -> False
    (Op_div _ _) -> False
    (Op_add _ _) -> False
```

```
    (Op_sub _ _) -> False
    _            -> True
isHigherPrecendence exp1@(Op_eq _ _) exp2 =
  case exp2 of
    (Op_not _  )  -> True
    (Op_and _ _)  -> True
    (Op_or  _ _)  -> True
    -- constants has lowest precendence
    (Lval _) -> True
    (BoolConst _) -> True
    (IntConst _) -> True
    (StrConst _) -> True
    _             -> False
isHigherPrecendence exp1@(Op_neq _ _) exp2 =
  case exp2 of
    (Op_not _  )  -> True
    (Op_and _ _)  -> True
    (Op_or  _ _)  -> True
    -- constants has lowest precendence
    (Lval _) -> True
    (BoolConst _) -> True
    (IntConst _) -> True
    (StrConst _) -> True
    _             -> False
isHigherPrecendence exp1@(Op_less _ _) exp2 =
  case exp2 of
    (Op_not _  )  -> True
    (Op_and _ _)  -> True
    (Op_or  _ _)  -> True
    -- constants has lowest precendence
    (Lval _) -> True
    (BoolConst _) -> True
    (IntConst _) -> True
    (StrConst _) -> True
    _             -> False
isHigherPrecendence exp1@(Op_less_eq _ _) exp2 =
  case exp2 of
    (Op_not _  )  -> True
    (Op_and _ _)  -> True
    (Op_or  _ _)  -> True
    -- constants has lowest precendence
    (Lval _) -> True
    (BoolConst _) -> True
    (IntConst _) -> True
    (StrConst _) -> True
    _             -> False
isHigherPrecendence exp1@(Op_large _ _) exp2 =
  case exp2 of
    (Op_not _  )  -> True
    (Op_and _ _)  -> True
    (Op_or  _ _)  -> True
    -- constants has lowest precendence
    (Lval _) -> True
    (BoolConst _) -> True
    (IntConst _) -> True
    (StrConst _) -> True
```

```
    _             -> False
isHigherPrecedence exp1@(Op_large_eq _ _) exp2 =
  case exp2 of
    (Op_not _  )  -> True
    (Op_and _ _)  -> True
    (Op_or  _ _)  -> True
    -- constants has lowest precendence
    (Lval _) -> True
    (BoolConst _) -> True
    (IntConst _) -> True
    (StrConst _) -> True
    _             -> False
isHigherPrecedence exp1@(Op_not _) exp2 =
  case exp2 of
    (Op_and _ _)  -> True
    (Op_or  _ _)  -> True
    -- constants has lowest precendence
    (Lval _) -> True
    (BoolConst _) -> True
    (IntConst _) -> True
    (StrConst _) -> True
    _             -> False
isHigherPrecedence exp1@(Op_and _ _) exp2 =
  case exp2 of
    (Op_or  _ _)  -> True
    -- constants has lowest precendence
    (Lval _) -> True
    (BoolConst _) -> True
    (IntConst _) -> True
    (StrConst _) -> True
    _             -> False
isHigherPrecedence exp1@(Op_or _ _) exp2
   -- constants has lowest precendence
  | (not (hasOperatorExp exp2)) = True
  | otherwise = False
isHigherPrecedence _ _ = False

-- exp1 has same precendence as exp2
isSamePrecedence :: Exp -> Exp -> Bool
isSamePrecedence exp1@(Op_neg _) exp2 =
  case exp2 of
    (Op_neg _  ) -> True
    _            -> False
isSamePrecedence exp1@(Op_mul _ _) exp2 =
  case exp2 of
    (Op_mul _ _) -> True
    (Op_div _ _) -> True
    _            -> False
isSamePrecedence exp1@(Op_div _ _) exp2 =
  case exp2 of
    (Op_mul _ _) -> True
    (Op_div _ _) -> True
    _            -> False
isSamePrecedence exp1@(Op_add _ _) exp2 =
  case exp2 of
    (Op_add _ _) -> True
```

```
    (Op_sub _ _) -> True
    _            -> False
isSamePrecendence exp1@(Op_sub _ _) exp2 =
  case exp2 of
    (Op_add _ _) -> True
    (Op_sub _ _) -> True
    _            -> False
isSamePrecendence exp1@(Op_eq _ _) exp2       = False -- relational
isSamePrecendence exp1@(Op_neq _ _) exp2      = False -- relational
isSamePrecendence exp1@(Op_less _ _) exp2     = False -- relational
isSamePrecendence exp1@(Op_less_eq _ _) exp2  = False -- relational
isSamePrecendence exp1@(Op_large _ _) exp2    = False -- relational
isSamePrecendence exp1@(Op_large_eq _ _) exp2 = False -- relational
isSamePrecendence exp1@(Op_not _) exp2 =
  case exp2 of
    (Op_not _)  -> True
    _           -> False
isSamePrecendence exp1@(Op_and _ _) exp2 =
  case exp2 of
    (Op_and _ _)  -> True
    _             -> False
isSamePrecendence exp1@(Op_or _ _) exp2 =
  case exp2 of
    (Op_or  _ _)  -> True
    _             -> False
isSamePrecendence _ _ = False

-- exp1 has smaller precendence than exp2 if it is not higher nor same
isSamllerPrecendence :: Exp -> Exp -> Bool
isSamllerPrecendence exp1 exp2 = (not (isHigherPrecendence exp1 exp2)) &&
                                 (not (isSamePrecendence   exp1 exp2))

-- does expression has operator in it?
hasOperatorExp :: Exp -> Bool
hasOperatorExp (Lval _) = False
hasOperatorExp (BoolConst _) = False
hasOperatorExp (IntConst _) = False
hasOperatorExp (StrConst _) = False
hasOperatorExp _ = True

-- return true if parent is sub/div and child has same precendence as parent
--    pexp: parent expression
--    cexp: child  expression
isDivSubParentSamePrecChild :: Exp -> Exp -> Bool
isDivSubParentSamePrecChild pexp@(Op_div _ _) cexp@(Op_div _ _) = True  -- / with a right child of / need
isDivSubParentSamePrecChild pexp@(Op_div _ _) cexp@(Op_mul _ _) = True  -- / with a right child of * need
isDivSubParentSamePrecChild pexp@(Op_sub _ _) cexp@(Op_sub _ _) = True  -- - (sub) with a right child of -
isDivSubParentSamePrecChild pexp@(Op_sub _ _) cexp@(Op_add _ _) = True  -- - (sub) with a right child of +
isDivSubParentSamePrecChild _ _ = False

-- True if Integer division happens
--    Integer division: 3 * (5 / 3) = 3 * 1 = 3, but (3 * 5) / 3 = 15 / 3 = 5,  so parens needed
isIntegerDision :: Exp -> Exp -> Bool
isIntegerDision pexp@(Op_mul _ _) cexp@(Op_div _ _) = True
isIntegerDision _ _ = False

```

```
276 -- some notation:
277 --    pexp: parent      expression (definitely has operator)
278 --    exp1: left child  expression
279 --    exp2: right child expression
280 -- turn binary expression's left child to string
281 strBinaryExpLChild :: Exp -> Exp -> String
282 strBinaryExpLChild pexp exp1
283   -- left child (with operator) has lower precendence suggests a parens
284   | (isSamllerPrecendence exp1 pexp) && (hasOperatorExp exp1) = surroundByParens (strExp exp1)
285   -- no parens
286   | otherwise = strExp exp1
287
288 -- True if expression is "not" <exp>
289 isNotExp :: Exp -> Bool
290 isNotExp (Op_not _) = True
291 isNotExp _ = False
292
293 -- some notation:
294 --    pexp: parent      expression (definitely has operator)
295 --    exp1: left child  expression
296 --    exp2: right child expression
297 -- turn binary expression's right child to string
298 strBinaryExpRChild :: Exp -> Exp -> String
299 strBinaryExpRChild pexp exp2
300   -- right child (with operator) has lower precendence (except "not" operator) or
301   --    same predence as parent (if parent is div(//) or sub(-)) or
302   --    integer division happens
303   -- suggests a parens
304   | (hasOperatorExp exp2) &&
305     ((isDivSubParentSamePrecChild pexp exp2) ||
306       (isSamllerPrecendence exp2 pexp) ||
307       (isIntegerDision pexp exp2)
308     ) &&
309     (not (isNotExp exp2))
310       = surroundByParens (strExp exp2)
311   -- no parens
312   | otherwise = strExp exp2
313
314 -- some notation:
315 --    pexp: parent      expression
316 --    exp1: left child  expression
317 --    exp2: right child expression
318 strExp :: Exp -> String
319 -- <lvalue>
320 strExp (Lval lValue) = strLValue lValue
321 -- <const>
322 strExp (BoolConst booleanLiteral) = strBooleanLiteral booleanLiteral
323 -- <const>
324 strExp (IntConst integerLiteral) = show integerLiteral
325 -- <const>
326 -- White space, and upper/lower case, should be preserved inside strings.
327 -- stringLiterals
328 strExp (StrConst stringLiteral) = "\"" ++ stringLiteral ++ "\""
329 -- <unop: "-"> <exp>
330 -- no space after unary minus
331 strExp pexp@(Op_neg exp)
```

```
  | (hasOperatorExp exp) && (not (isSamePrecendence pexp exp)) = "-" ++ surroundByParens (strExp exp) -- n
  | otherwise         = "-" ++ (strExp exp) -- no need to parens constants/Lval
-- <exp> <binop: "*"> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_mul exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " * " ++ (strBinaryExpRChild pexp exp2)
-- <exp> <binop: "/"> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_div exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " / " ++ (strBinaryExpRChild pexp exp2)
-- <exp> <binop: "+"> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_add exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " + " ++ (strBinaryExpRChild pexp exp2)
-- <exp> <binop: "-"> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_sub exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " - " ++ (strBinaryExpRChild pexp exp2)
-- <exp> <binop: "="> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_eq exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " = " ++ (strBinaryExpRChild pexp exp2)
-- <exp> <binop: "!="> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_neq exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " != " ++ (strBinaryExpRChild pexp exp2
-- <exp> <binop: "<"> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_less exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " < " ++ (strBinaryExpRChild pexp exp2
-- <exp> <binop: "<="> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_less_eq exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " <= " ++ (strBinaryExpRChild pexp
-- <exp> <binop: ">"> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_large exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " > " ++ (strBinaryExpRChild pexp exp
-- <exp> <binop: ">="> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_large_eq exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " >= " ++ (strBinaryExpRChild pexp
-- <unop: not> <exp>
-- There should be a single space after not.
strExp pexp@(Op_not exp)
  -- need to parens expression with operator and (child's precendence is smaller)
  | (hasOperatorExp exp) && (isSamllerPrecendence exp pexp) = "not " ++ surroundByParens (strExp exp)
  | otherwise         = "not " ++ (strExp exp) -- no need to parens constants/Lval
-- <exp> <binop: and> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_and exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " and " ++ (strBinaryExpRChild pexp exp
-- <exp> <binop: or> <exp>
-- Single space should surround 12 binary operators.
strExp pexp@(Op_or exp1 exp2) = (strBinaryExpLChild pexp exp1) ++ " or " ++ (strBinaryExpRChild pexp exp2)


-- Int: indentation level
strStmt :: Int -> Stmt -> String
-- In a procedure body, each statement should start on a new line. So ++ newline in each's end
strStmt indentLevel (Assign lValue exp) =
  -- <lvalue> <- <exp>;
  -- Single spaces should surround the assignment operator <-
  (addIndentation indentLevel) ++ (strLValue lValue) ++ " <- " ++ (strExp exp) ++ semicolon ++ newline
strStmt indentLevel (Read lValue) =
  -- read <lvalue>;
  (addIndentation indentLevel) ++ "read " ++ (strLValue lValue) ++ semicolon ++ newline
```

```
strStmt indentLevel (Write exp) =
  -- write <exp>;
  (addIndentation indentLevel) ++ "write " ++ (strExp exp) ++ semicolon ++ newline
strStmt indentLevel (Writeln exp) =
  -- writeln <exp>;
  (addIndentation indentLevel) ++ "writeln " ++ (strExp exp) ++ semicolon ++ newline
strStmt indentLevel (Call ident exps) =
  -- call <id>(<exp-list>);
  --     where <exp-list> is a (possibly empty according to parser) comma-separated list of expressions.
  (addIndentation indentLevel) ++ "call " ++ ident ++ surroundByParens (intercalate comma (map strExp exps
-- thenStmts is non-empty according to parser, elseStmts is possible empty according to parser
strStmt indentLevel (If exp thenStmts elseStmts) =
  -- IF elseStmts is empty: according to parser: if <exp> then <stmt-list> fi
  if null elseStmts then
    -- "if ... then" should be printed on one line, irrespective of the size of the intervening expression
    (addIndentation indentLevel) ++ "if " ++ (strExp exp) ++ " then" ++ newline ++
    -- more indentation
    (concatMap (strStmt (indentLevel+1)) thenStmts) ++
    -- the terminating fi should be indented exactly as the corresponding if.
    (addIndentation indentLevel) ++ "fi" ++ newline
  -- OTHERWISE            : according to parser: if <expr> then <stmt-list> else <stmt-list> fi
  else
    -- "if ... then" should be printed on one line, irrespective of the size of the intervening expression
    (addIndentation indentLevel) ++ "if " ++ (strExp exp) ++ " then" ++ newline ++
    -- more indentation
    (concatMap (strStmt (indentLevel+1)) thenStmts) ++
    (addIndentation indentLevel) ++ "else" ++ newline ++
    -- more indentation
    (concatMap (strStmt (indentLevel+1)) elseStmts) ++
    -- the terminating fi and else should be indented exactly as the corresponding if.
    (addIndentation indentLevel) ++ "fi" ++ newline
-- stmts is non-empty according to parser
strStmt indentLevel (While exp stmts) =
  -- In a while statement, "while ... do" should be printed on one line,
  --    irrespective of the size of the intervening expression
  (addIndentation indentLevel) ++ "while " ++ (strExp exp) ++ " do" ++ newline ++
  -- more indentation
  (concatMap (strStmt (indentLevel+1)) stmts) ++
  -- The terminating od should be indented exactly as the corresponding while.
  (addIndentation indentLevel) ++ "od" ++ newline

-----------------------------------------------------------------
--  Record toString function
-----------------------------------------------------------------
-- field declaration is of:
--   1. boolean or integer
--   2. followed by an identifier (the field name).
strFieldDecl :: FieldDecl -> String
strFieldDecl (FieldDecl baseType fieldName) = (strBaseType baseType) ++ " " ++ fieldName

-- non-empty input list fieldDecls@(x:xs) according to parser
strFieldDecls :: [FieldDecl] -> String
-- first field decl starts with {
strFieldDecls (x:xs) = (addIndentation 1) ++ "{ " ++ (strFieldDecl x) ++ newline ++
-- rest start with ;
                       (concatMap
```

```
                          (\y -> (addIndentation 1) ++ "; " ++ (strFieldDecl y) ++ newline)
                          xs
                      )

-- convert record to string
-- A record type definition involving n fields should be printed on n + 2 lines,
-- as follows:
--     1. The first line contains the word record.
--     2. The remaining lines should be indented, with the first n containing one field declaration each
--        (the first preceded by a left brace and a single space, the rest preceded
--          by a semicolon and a single space),
--          see above strFieldDecls
--     3. and with the last line containing the record name, preceded by a right
--          brace and a single space, and followed by a semicolon;
strRecord :: Record -> String
strRecord (Record fieldDecls recordName) =
  "record" ++ newline ++
  (strFieldDecls fieldDecls) ++
  (addIndentation 1) ++ "} " ++ recordName ++ semicolon ++ newline


-----------------------------------------------------------------
--  Array toString function
-----------------------------------------------------------------
-- An array type definition should be printed on a single line.
-- It contains the word array,
-- followed by a positive integer in square brackets all without intervening
--      white space.
-- That string, the type, and the type alias, should be separated by single
--    spaces, and the whole line terminated by a semicolon.
strArray :: Array -> String
strArray (Array arraySize arrayType arrayName) =
  "array" ++ surroundByBrackets (show arraySize) ++ " " ++ (strDataType arrayType) ++ " " ++ arrayName ++ s


-----------------------------------------------------------------
--  Procedure toString functions
-----------------------------------------------------------------
strParameter :: Parameter -> String
strParameter (DataParameter dataType paraName) = (strDataType dataType) ++ " " ++ paraName
strParameter (BooleanVal paraName) = "boolean val " ++ paraName
strParameter (IntegerVal paraName) = "integer val " ++ paraName

-- The procedure head (that is, the keyword, procedure name, and list of formal
--     parameters) should be on a single line.
strProcedureHeader :: ProcedureHeader -> String
strProcedureHeader (ProcedureHeader procedureName parameters) =
  procedureName ++ " " ++ (surroundByParens (intercalate comma (map strParameter parameters)))

-- A variable declaration consists of
--   a) a type name (boolean, integer, or a type alias),
--   b) followed by a 1+ comma-separated list of
--        identifiers,
--     i)  the list terminated with a semicolon.
--     ii) There may be any number of variable declarations, in any order.
strVariableDecl :: VariableDecl -> String
```

```
strVariableDecl (VariableDecl dataType varNames) =
  -- Within each procedure, declarations and top-level statements should be indented.
  (addIndentation 1) ++ (strDataType dataType) ++ " " ++ (intercalate comma varNames)  ++ semicolon ++
  -- Each variable declaration should be on a separate line.
  newline

-- variableDecls can be empty according to parser
-- stmts is non-empty according to parser
strProcedureBody :: ProcedureBody -> String
strProcedureBody (ProcedureBody variableDecls stmts) =
  (concatMap strVariableDecl variableDecls) ++
  -- The { and } that surround a procedure body should begin at the start of a
  --    line (no indentation).
  -- Moreover, these delimiters should appear alone, each making up a single line.
  "{" ++ newline ++
  -- Within each procedure, declarations and top-level statements should be indented.
  concatMap (strStmt 1) stmts ++
  "}" ++ newline

-- convert procedure to string
strProcedure :: Procedure -> String
-- The keyword procedure should begin at the start of a line (no indentation)
-- The procedure head (that is, the keyword, procedure name, and list of formal
--     parameters) should be on a single line.
strProcedure (Procedure ph pb) =
  "procedure " ++ (strProcedureHeader ph) ++ newline ++
  (strProcedureBody pb)


-----------------------------------------------------------------
--  Program toString function
-----------------------------------------------------------------
strProgram :: Program -> String
-- If there are no record and array type definitions, the first procedure should
--    start on line 1.
strProgram (Program [] [] procedures) = intercalate newline (map strProcedure procedures)
-- Otherwise there should be a single blank line between the type definitions
--    and the first procedure.
strProgram (Program records arraies procedures) =
  -- Each type definition should start on a new line, and there should be no
  --    blank lines between type definitions. So below two has no newline in between
  concatMap strRecord records ++
  concatMap strArray arraies ++
  newline ++
  -- Consecutive procedure definitions should be separated by a single blank line.
  intercalate newline (map strProcedure procedures)
```

Roo.hs

```
1  ------------------------------------------------------------
2  -- COMP90045 Programming Language Implementation Project --
3  --                      Roo Compiler                      --
4  --  Implemented by Xulin Yang                            --
5  ------------------------------------------------------------
6  module Main (main)
7  where
8  import RooParser (ast)
9  import PrettyRoo (pp)
10 import System.Environment (getProgName, getArgs)
11 import System.Exit (exitWith, ExitCode(..))
12
13 data Task
14   = Parse | Pprint
15     deriving (Eq, Show)
16
17 main :: IO ()
18 main
19   = do
20       progname <- getProgName
21       args <- getArgs
22       task <- checkArgs progname args
23       case task of
24         Parse
25           -> do
26                let [_, filename] = args
27                input <- readFile filename
28                let output = ast input
29                case output of
30                  Right tree
31                    -> putStrLn (show tree)
32                  Left err
33                    -> do putStrLn "Parse error at "
34                          print err
35                          exitWith (ExitFailure 2)
36         Pprint
37           -> do
38                let [_, filename] = args
39                input <- readFile filename
40                let output = ast input
41                case output of
42                  Right tree
43                    -> putStr (pp tree)
44                  Left err
45                    -> do putStrLn "Parse error at "
46                          print err
47                          exitWith (ExitFailure 2)
48
49 checkArgs :: String -> [String] -> IO Task
50 checkArgs _ ['-':_]
51   = do
```

```
52        putStrLn ("Missing filename")
53        exitWith (ExitFailure 1)
54  checkArgs _ ["-a", filename]
55    = return Parse
56  checkArgs _ ["-p", filename]
57    = return Pprint
58  checkArgs progname _
59    = do
60        putStrLn ("Usage: " ++ progname ++ " [-p] filename")
61        exitWith (ExitFailure 1)
62
```

RooAST.hs

```
1  ------------------------------------------------------------
2  -- COMP90045 Programming Language Implementation Project --
3  --                    Roo Compiler                        --
4  --  Implemented by Xulin Yang                             --
5  --  read from the bottom to top                           --
6  ------------------------------------------------------------
7  module RooAST where
8
9  ----------------------------------
10 -- Terminology:
11 -- 0+: zero or more/possible empty
12 -- 1+: one or more/ non empty
13 --     both 0+, 1+ are stored in list [] but 1+ will be implemented in parser not here
14 ----------------------------------
15
16 ----------------------------------
17 -- Specification of an AST for Roo
18 ----------------------------------
19
20 -- Identifier: String
21 type Ident = String
22
23 -- Base type: boolean, integer type indicator
24 --     Not necessary to have string as no variable/parameter/declaration has string type
25 data BaseType
26   = BooleanType
27   | IntegerType
28     deriving (Show, Eq)
29
30 -- A boolean literal is false or true.
31 type BooleanLiteral = Bool
32 -- An integer literal is a sequence of digits, only stores natural number in our parser implementation
33 type IntegerLiteral = Int
34 -- A string literal is a sequence of characters between double quotes.
35 --   The sequence itself cannot contain double quotes or newline/tab characters.
36 --   It may, however, contain '" ', '\n', and '\t', respectively, to represent
37 --   those characters.
38 type StringLiteral = String
39
40 -- User custermized record type, stored as string
41 type AliasType = String
42
43 -- for Array, VariableDecl: they have either boolean, integer, or a type alias data type
44 --     factored out for reuse purpose
45 data DataType
46   = BasyDataType BaseType
47   | AliasDataType AliasType
48     deriving (Show, Eq)
49
50 -- An lvalue (<lvalue>) has four (and only four) possible forms:
51 --     An example lvalue is point[0].xCoord
```

```
data LValue
  = LId Ident                       -- <id>
  | LDot Ident Ident                -- <id>.<id>
  | LBrackets Ident Exp             -- <id>[<exp>]
  | LBracketsDot Ident Exp Ident    -- <id>[<exp>].<id>
    deriving (Show, Eq)

-- expression operators:
--     All the operators on the same line have the same precedence,
--         and the ones on later lines have lower precedence;
--     The six relational operators are non-associative
--         so, for example, a = b = c is not a well-formed expression).
--     The six remaining binary operators are left-associative.
-- -              |unary            |
-- * /            |binary and infix |left-associative
-- + -            |binary and infix |left-associative
-- = != < <= > >= |binary and infix |relational, non-associative
-- not            |unary            |
-- and            |binary and infix |left-associative
-- or             |binary and infix |left-associative
data Exp
  = Lval LValue                 -- <lvalue>
  | BoolConst BooleanLiteral    -- <const> where <const> is the syntactic category of boolean, integer, and
  | IntConst IntegerLiteral
  | StrConst StringLiteral
                                -- ( <exp> ) is ignored here but handelled in parser
  | Op_or Exp Exp               -- <exp> <binop: or> <exp>
  | Op_and Exp Exp              -- <exp> <binop: and> <exp>
  | Op_eq  Exp Exp              -- <exp> <binop: "="> <exp>
  | Op_neq  Exp Exp             -- <exp> <binop: "!="> <exp>
  | Op_less  Exp Exp            -- <exp> <binop: "<"> <exp>
  | Op_less_eq  Exp Exp         -- <exp> <binop: "<="> <exp>
  | Op_large  Exp Exp           -- <exp> <binop: ">"> <exp>
  | Op_large_eq  Exp Exp        -- <exp> <binop: ">="> <exp>
  | Op_add  Exp Exp             -- <exp> <binop: "+"> <exp>
  | Op_sub  Exp Exp             -- <exp> <binop: "-"> <exp>
  | Op_mul  Exp Exp             -- <exp> <binop: "*"> <exp>
  | Op_div  Exp Exp             -- <exp> <binop: "/"> <exp>
  | Op_not Exp                  -- <unop: not> <exp>
  | Op_neg Exp                  -- <unop: "-"> <exp>
    deriving (Show, Eq)

-- Stmt has following two category:
--  1) atom statement:
--      <lvalue> <- <exp> ;
--      read <lvalue> ;
--      write <exp> ;
--      writeln <exp> ;
--      call <id>(<exp-list>) ;
--          where <exp-list> is a 0+ comma-separated list of expressions.
--  2) composite statement:
--      if <exp> then <stmt-list> else <stmt-list> fi
--      if <exp> then <stmt-list> fi # just make above second [Stmt] empty
--      while <exp> do <stmt-list> od
--          where <stmt-list> is a 1+ sequence of statements, atomic or composite
--
```

```
-- the data structure for above grammer are given accordingly below
data Stmt
  -- 1) atom statement:
  = Assign LValue Exp
  | Read LValue
  | Write Exp
  | Writeln Exp
  | Call Ident [Exp]
  -- 2) composite statement:
  | If Exp [Stmt] [Stmt]
  | While Exp [Stmt]
    deriving (Show, Eq)

-- Each formal parameter has two components (in the given order):
--  1. a parameter type/mode indicator, which is one of these five:
--    a) a type alias,
--    b) boolean,
--    c) integer,
--    d) boolean val
--    e) integer val
--  2. an identifier
data Parameter
  = DataParameter DataType Ident -- a) b) c) above
  | BooleanVal Ident             -- d)       above
  | IntegerVal Ident             -- e)       above
    deriving (Show, Eq)

-- The header has two components (in this order):
--   1. an identifier (the procedure's name), and
--   2. a comma-separated list of 0+ formal parameters within a pair
--        of parentheses (so the parentheses are always present).
data ProcedureHeader
  = ProcedureHeader Ident [Parameter]
    deriving (Show, Eq)

-- A variable declaration consists of
--   a) a type name (boolean, integer, or a type alias),
--   b) followed by a 1+ comma-separated list of
--        identifiers,
--     i)  the list terminated with a semicolon.
--     ii) There may be any number of variable declarations, in any order.
data VariableDecl
  = VariableDecl DataType [Ident]
    deriving (Show, Eq)

-- procedure body consists of 0+ local variable declarations,
--   1. A variable declaration consists of
--     a) a type name (boolean, integer, or a type alias),
--     b) followed by a 1+ comma-separated list of identifiers,
--       i)  the list terminated with a semicolon.
--       ii) There may be any number of variable declarations, in any order.
--   2. followed by a 1+ sequence of statements,
data ProcedureBody
  = ProcedureBody [VariableDecl] [Stmt]
    deriving (Show, Eq)
```

```haskell
-- Each procedure consists of (in the given order):
--   1. the keyword procedure,
--   2. a procedure header, and
--   3. a procedure body.
data Procedure
  = Procedure ProcedureHeader ProcedureBody
    deriving (Show, Eq)

-- array type definition consists of (in the given order):
--   1. the keyword array,
--   2. a (positive) integer literal enclosed in square brackets,
--   3. a type name which is either an identifier (a type alias) or one of
--      boolean and integer,
--   4. an identifier (giving a name to the array type), and
--   5. a semicolon.
data Array
  = Array IntegerLiteral DataType Ident
    deriving (Show, Eq)

-- field declaration is of:
--   1. boolean or integer
--   2. followed by an identifier (the field name).
data FieldDecl
  = FieldDecl BaseType Ident
    deriving (Show, Eq)

-- record consists of:
--   1. the keyword record,
--   2. a 1+ list of field declarations, separated by semicolons,
--        the whole list enclosed in braces,
--   3. an identifier, and
--   4. a semicolon.
data Record
  = Record [FieldDecl] Ident
    deriving (Show, Eq)

-- A Roo program consists of
--   1. 0+ record type definitions, followed by
--   2. 0+ array type definitions, followed by
--   3. 1+ procedure definitions.
data Program
  = Program [Record] [Array] [Procedure]
    deriving (Show, Eq)
```

==========

RooParser.hs

==========

```
1  ------------------------------------------------------------
2  -- COMP90045 Programming Language Implementation Project --
3  --                    Roo Compiler                        --
4  --  Implemented by Xulin Yang                             --
5  --  read from the bottom to top                           --
6  ------------------------------------------------------------
7  module RooParser (ast)
8  where
9  import RooAST
10 import Text.Parsec
11 import Text.Parsec.Language (emptyDef)
12 import Text.Parsec.Expr
13 import qualified Text.Parsec.Token as Q
14 import System.Environment
15 import System.Exit
16 import Debug.Trace (trace)
17
18 type Parser a
19    = Parsec String Int a
20
21 scanner :: Q.TokenParser Int
22 scanner
23    = Q.makeTokenParser
24      (emptyDef
25      { Q.commentLine     = "#"
26      , Q.nestedComments  = True
27      , Q.identStart      = letter
28      -- An identifir is a non-empty sequence of alphanumeric characters,
29      --  underscore and apostrophe ('), and it must start with a (lower or upper case) letter.
30      , Q.identLetter     = alphaNum <|> char '_' <|> char '\''
31      , Q.opStart         = oneOf "+-*<"
32      , Q.opLetter        = oneOf "="
33      , Q.reservedNames   = joeyReserved
34      , Q.reservedOpNames = joeyOpnames
35      })
36
37 whiteSpace    = Q.whiteSpace scanner
38 lexeme        = Q.lexeme scanner
39 natural       = Q.natural scanner
40 identifier    = Q.identifier scanner
41 semi          = Q.semi scanner
42 comma         = Q.comma scanner
43 dot           = Q.dot scanner
44 parens        = Q.parens scanner
45 braces        = Q.braces scanner
46 brackets      = Q.brackets scanner
47 squares       = Q.squares scanner
48 reserved      = Q.reserved scanner
49 reservedOp    = Q.reservedOp scanner
50 stringLiteral = Q.stringLiteral scanner
51
```

```
52  joeyReserved, joeyOpnames :: [String]
53
54  -- reserved words according to the specification
55  joeyReserved
56    = ["and", "array", "boolean", "call", "do", "else", "false", "fi", "if",
57       "integer", "not", "od", "or", "procedure", "read", "record", "then",
58       "true", "val", "while", "write", "writeln"]
59
60  -- reserved operators from specification
61  -- 12 binary oprator (and, or; above reserved string);
62  -- 2 unary: not (above reserved string), -;
63  -- assignment operator <-
64  joeyOpnames
65    = [ "+", "-", "*", "/", "=", "!=", "<", "<=", ">", ">=", "<-"]
66
67  ---------------------------------------------------------------
68  --  Note: 0+ is ensured using many/sepBy and 1+ using many1/sepBy1
69  ---------------------------------------------------------------
70
71  ---------------------------------------------------------------
72  --  parse reused base type integer/boolean; no string here as mentioned in AST
73  --  parse reused data type integer/boolean/type alias
74  ---------------------------------------------------------------
75  pBaseType :: Parser BaseType
76  pBaseType
77    = do
78        reserved "boolean"
79        return BooleanType
80      <|>
81      do
82        reserved "integer"
83        return IntegerType
84      <?>
85        "base type"
86
87  pDataType :: Parser DataType
88  pDataType
89    =
90      do
91        baseType <- pBaseType
92        return (BasyDataType baseType)
93      <|>
94      do
95        alias <- identifier
96        return (AliasDataType alias)
97      <?>
98        "data type"
99
100
101 ---------------------------------------------------------------
102 --  parse literals
103 ---------------------------------------------------------------
104 pBooleanLiteral :: Parser BooleanLiteral
105 pBooleanLiteral
106  =
107    do { reserved "true"; return (True) }
```

```
    <|>
    do { reserved "false"; return (False) }
    <?>
       "boolean literal"

pIntegerLiteral :: Parser IntegerLiteral
pIntegerLiteral
  =
    do
      n <- natural <?> "number"
      return (fromInteger n :: Int)
    <?>
      "Integer Literal"

-- don't accept newline, tab, quote but "\n", "\t", "\"" <- two character string should still be accepted
pcharacter :: Parser String
pcharacter
  =
    try(
      do
        string ('\\':['n'])
        return (['\\', 'n'])
    )
    <|>
    try(
      do
        string ('\\':['t'])
        return (['\\', 't'])
    )
    <|>
    try(
      do
        string ('\\':['"'])
        return (['\\', '"'])
    )
    <|>
    do
      c <- noneOf ['\n', '\t', '"']
      return ([c])
    <?>
      "any character except newline, tab, quote"

-- Parser for string
pString :: Parser String
pString
  =
    do
      -- String is surrounded by two quotes
      char '"'
      -- Parse characters except newline / tab characters and quotes
      str <- many pcharacter
      char '"' <?> "\'\"\' to wrap the string"
      spaces -- consumes following spaces
      return (concat str)
    <?>
      "string cannot has newline, quote, tab"
```

```
164
165 pStringLiteral :: Parser StringLiteral
166 pStringLiteral
167   =
168     do
169       s <- pString
170       return (s)
171     <?>
172       "string literal"
173
174
175 -----------------------------------------------------------------
176 -- An lvalue (<lvalue>) has four (and only four) possible forms:
177 --     <id>
178 --     <id>.<id>
179 --     <id>[<exp>]
180 --     <id>[<exp>].<id>
181 -- An example lvalue is point[0].xCoord
182 -----------------------------------------------------------------
183 pLValue :: Parser LValue
184 pLValue
185   = try (
186       do
187         ident1 <- identifier
188         exp <- brackets pExp
189         dot
190         ident2 <- identifier
191         return (LBracketsDot ident1 exp ident2)
192    )
193     <|>
194     try (
195       do
196         ident <- identifier
197         exp <- brackets pExp
198         return (LBrackets ident exp)
199     )
200     <|>
201     try (
202       do
203         ident1 <- identifier
204         dot
205         ident2 <- identifier
206         return (LDot ident1 ident2)
207     )
208     <|>
209     do
210       ident <- identifier
211       return (LId ident)
212     <?>
213       "LValue"
214
215 -----------------------------------------------------------------
216 --  pExp is the main parser for expression.
217 --
218 --  It is built using Parces's powerful
219 --  buildExpressionParser and takes into account the operator
```

```
220 --  precedences and associativity specified in 'opTable' below.
221 ---------------------------------------------------------------------
222 prefix name fun
223   = Prefix (do { reservedOp name
224                ; return fun
225                }
226            )
227
228 binary name op
229   = Infix (do { reservedOp name
230               ; return op
231               }
232           ) AssocLeft
233
234 relation name rel
235   = Infix (do { reservedOp name
236               ; return rel
237               }
238           ) AssocNone
239
240 -- expression operators:
241 --     All the operators on the same line have the same precedence,
242 --         and the ones on later lines have lower precedence;
243 --     The six relational operators are non-associative
244 --         so, for example, a = b = c is not a well-formed expression).
245 --     The six remaining binary operators are left-associative.
246 -- -               |unary            |
247 -- * /             |binary and infix |left-associative
248 -- + -             |binary and infix |left-associative
249 -- = != < <= > >= |binary and infix |relational, non-associative
250 -- not             |unary            |
251 -- and             |binary and infix |left-associative
252 -- or              |binary and infix |left-associative
253 opTable
254   = [ [ prefix   "-"   Op_neg    ]
255     , [ binary   "*"   Op_mul    , binary   "/"  Op_div  ]
256     , [ binary   "+"   Op_add    , binary   "-"  Op_sub  ]
257     , [ relation "="   Op_eq     , relation "!=" Op_neq  , relation "<"  Op_less
258       , relation "<="  Op_less_eq, relation ">"  Op_large, relation ">=" Op_large_eq ]
259     , [ prefix   "not" Op_not    ]
260     , [ binary   "and" Op_and    ]
261     , [ binary   "or"  Op_or     ]
262     ]
263
264 pExp :: Parser Exp
265 pExp
266   = buildExpressionParser opTable pFac
267     <?>
268       "expression"
269
270 pFac :: Parser Exp
271 pFac
272   = choice [parens pExp, -- ( <exp> )
273             pLval,
274             pBoolConst,
275             pIntConst,
```

```
276             pStrConst,
277             pNeg         -- used to parse expression like ------1 (arbitrary unary minus before expression
278            ]
279     <?>
280       "simple expression"
281
282 pLval, pBoolConst, pIntConst, pStrConst :: Parser Exp
283 pLval
284   =
285     do
286       lval <- pLValue
287       return (Lval lval)
288     <?>
289       "lval expression"
290
291 pBoolConst
292   =
293     do
294       b <- pBooleanLiteral
295       return (BoolConst b)
296     <?>
297       "bool const/literal expression"
298
299 pIntConst
300   =
301     do
302       i <- pIntegerLiteral
303       return (IntConst i)
304     <?>
305       "int const/literal expression"
306
307 pStrConst
308   =
309     do
310       s <- pStringLiteral
311       return (StrConst s)
312     <?>
313       "string const/literal expression"
314
315 -- parse arbitrary unary minus in pFac
316 pNeg :: Parser Exp
317 pNeg
318   =
319   do
320     reservedOp "-"
321     exp <- pExp
322     return (Op_neg exp)
323   <?>
324     "unary minus"
325
326
327 ---------------------------------------------------------------------
328 --  pStmt is the main parser for statements.
329 --  Statement related parsers
330 ---------------------------------------------------------------------
331 pStmt, pStmtAtom, pStmtComp :: Parser Stmt
```

```
332 -- atom statement:
333 --     <lvalue> <- <exp> ;
334 --     read <lvalue> ;
335 --     write <exp> ;
336 --     writeln <exp> ;
337 --     call <id> ( <exp-list> ) ;
338 --         where <exp-list> is a (possibly empty) comma-separated list of expressions.
339 -- composite statement:
340 --     if <expr> then <stmt-list> else <stmt-list> fi
341 --     if <exp> then <stmt-list> fi # just make second [Stmt] empty
342 --     while <expr> do <stmt-list> od
343 --         where <stmt-list> is a non-empty sequence of statements, atomic or composite
344 pStmt = choice [pStmtAtom, pStmtComp]
345
346 pStmtAtom
347   =
348     do
349       r <- choice [pAsg, pRead, pWrite, pWriteln, pCall]
350       -- all atomic stmt's semicolon is comsumed here
351       semi
352       return r
353     <?>
354       "atomic statement"
355
356 pAsg, pRead, pWrite, pWriteln, pCall :: Parser Stmt
357
358 -- parse: <lvalue> <- <exp> ;
359 pAsg
360   =
361     do
362       lvalue <- pLValue
363       reservedOp "<-"
364       rvalue <- pExp
365       return (Assign lvalue rvalue)
366     <?>
367       "assign"
368
369 -- parse: read <lvalue> ;
370 pRead
371   = do
372       reserved "read"
373       lvalue <- pLValue
374       return (Read lvalue)
375     <?>
376       "read"
377
378 -- parse: write <exp> ;
379 pWrite
380   = do
381       reserved "write"
382       expr <- pExp
383       return (Write expr)
384     <?>
385       "write"
386
387 -- parse: writeln <exp> ;
```

```
pWriteln
  = do
      reserved "writeln"
      expr <- pExp
      return (Writeln expr)
    <?>
      "writeln"

-- parse: call <id>(<exp-list>) ;
pCall
  = do
      reserved "call"
      ident <- identifier
      exprs <- parens (pExp `sepBy` comma) -- 0+ comma-separated list of expressions
      return (Call ident exprs)
    <?>
      "call"

pStmtComp = (choice [pIf, pWhile]) <?> "composite statement"

pIf, pWhile :: Parser Stmt

-- parse:
--    if <exp> then <stmt-list> else <stmt-list> fi
--    if <exp> then <stmt-list> fi # just make above second [Stmt] empty
pIf
  = do

      reserved "if"
      exp <- pExp
      reserved "then"
      stmts <- many1 pStmt
      -- check if there is an else statment
      -- if not, return empty
      estmts <- (
        do
          reserved "fi"
          return []
        <|>
        do
          reserved "else"
          -- else body can not be empty
          s <- many1 pStmt
          reserved "fi"
          return s
        )
      return (If exp stmts estmts)
    <?>
      "if"

-- parse: while <exp> do <stmt-list> od
pWhile
  = do
      reserved "while"
      exp <- pExp
      reserved "do"
```

```
444       stmts <- many1 pStmt -- a 1+ sequence of statements, atomic or composite
445       reserved "od"
446       return (While exp stmts)
447     <?>
448       "while"
449
450 ----------------------------------------------------------------
451 --  Procedure related parser
452 ----------------------------------------------------------------
453 -- Each formal parameter has two components (in the given order):
454 --  1. a parameter type/mode indicator, which is one of these five:
455 --    a) a type alias,
456 --    b) boolean,
457 --    c) integer,
458 --    d) boolean val
459 --    e) integer val
460 --  2. an identifier
461 pParameter :: Parser Parameter
462 pParameter
463   =
464     try(
465       do
466         -- parse boolean val variable
467         reserved "boolean"
468         reserved "val"
469         name <- identifier
470         return (BooleanVal name)
471     )
472     <|>
473     try(
474       do
475         -- parse integer val variable
476         reserved "integer"
477         reserved "val"
478         name <- identifier
479         return (IntegerVal name)
480     )
481     <|>
482     do
483       -- parse boolean/integer/type_alias variable
484       paraType <- pDataType
485       name <- identifier
486       return (DataParameter paraType name)
487     <?>
488       "parameter"
489
490 -- The header has two components (in this order):
491 --   1. an identifier (the procedure's name), and
492 --   2. a comma-separated list of 0+ formal parameters within a pair
493 --        of parentheses (so the parentheses are always present).
494 pProcedureHeader :: Parser ProcedureHeader
495 pProcedureHeader
496   =
497     do
498       procedureName <- identifier
499       parameters <- parens (pParameter `sepBy` comma)
```

```
      return (ProcedureHeader procedureName parameters)
    <?>
      "procedure header"

-- A variable declaration consists of
--   a) a type name (boolean, integer, or a type alias),
--   b) followed by a 1+ comma-separated list of
--        identifiers,
--     i)  the list terminated with a semicolon.
--     ii) There may be any number of variable declarations, in any order.
pVariable :: Parser VariableDecl
pVariable
  =
    do
      varType <- pDataType
      varNames <- (identifier `sepBy1` comma)
      semi
      return (VariableDecl varType varNames)
    <?>
      "variable"


-- procedure body consists of 0+ local variable declarations,
--   1. A variable declaration consists of
--     a) a type name (boolean, integer, or a type alias),
--     b) followed by a 1+ comma-separated list of identifiers,
--       i)  the list terminated with a semicolon.
--       ii) There may be any number of variable declarations, in any order.
--   2. followed by a 1+ sequence of statements,
pProcedureBody :: Parser ProcedureBody
pProcedureBody
  =
    do
      vars <- many pVariable
      stmts <- braces (many1 pStmt)
      return (ProcedureBody vars stmts)
    <?>
      "procedure body"

-- Each procedure consists of (in the given order):
--   1. the keyword procedure,
--   2. a procedure header, and
--   3. a procedure body.
pProcedure :: Parser Procedure
pProcedure
  =
    do
      reserved "procedure"
      procedureHeader <- pProcedureHeader
      procedureBody <- pProcedureBody
      return (Procedure procedureHeader procedureBody)
    <?>
      "procedure"

-----------------------------------------------------------------
--  Array related parser
```

```
556 ----------------------------------------------------------------
557 -- array type definition consists of (in the given order):
558 --   1. the keyword array,
559 --   2. a (positive) integer literal enclosed in square brackets,
560 --   3. a type name which is either an identifier (a type alias) or one of
561 --      boolean and integer,
562 --   4. an identifier (giving a name to the array type), and
563 --   5. a semicolon.
564 pArray :: Parser Array
565 pArray
566   =
567     do
568       reserved "array"
569       pos <- getPosition
570       arraySize <- brackets pIntegerLiteral
571       -- need to check arraySize > 0 (positive integer)
572       if arraySize == 0
573       then
574         error ("array size sould not be 0 at line: " ++ (show (sourceLine pos)) ++ ", column: " ++ (show (
575       else do
576         arrayType <- pDataType
577         arrayName <- identifier
578         semi
579         return (Array arraySize arrayType arrayName)
580       <?>
581         "array"
582
583 ----------------------------------------------------------------
584 --  Record related parser
585 ----------------------------------------------------------------
586 -- field declaration is of:
587 --   1. boolean or integer
588 --   2. followed by an identifier (the field name).
589 pFieldDecl :: Parser FieldDecl
590 pFieldDecl
591   =
592     do
593       fieldType <- pBaseType
594       fieldName <- identifier
595       return (FieldDecl fieldType fieldName)
596     <?>
597       "field declaration"
598
599 -- record consists of:
600 --   1. the keyword record,
601 --   2. a 1+ list of field declarations, separated by semicolons,
602 --        the whole list enclosed in braces,
603 --   3. an identifier, and
604 --   4. a semicolon.
605 pRecord :: Parser Record
606 pRecord
607   =
608     do
609       reserved "record"
610       recordFieldDecls <- braces (pFieldDecl `sepBy1` semi)
611       recordName <- identifier
```

```
612       semi
613       return (Record recordFieldDecls recordName)
614     <?>
615       "record"
616
617 ----------------------------------------------------------------
618 --  Program related parser
619 ----------------------------------------------------------------
620 -- A Roo program consists of
621 --   1. 0+ record type definitions, followed by
622 --   2. 0+ array type definitions, followed by
623 --   3. 1+ procedure definitions.
624 pProgram :: Parser Program
625 pProgram
626   = do
627       records <- many pRecord
628       arraies <- many pArray
629       procedures <- many1 pProcedure
630       return (Program records arraies procedures)
631     <?>
632       "program"
633
634
635 ----------------------------------------------------------------
636 -- main (given skeleton code)
637 ----------------------------------------------------------------
638
639 joeyParse :: Parser Program
640 joeyParse
641   = do
642       whiteSpace
643       p <- pProgram
644       eof
645       return p
646
647 ast :: String -> Either ParseError Program
648 ast input
649   =  runParser joeyParse 0 "" input
650
651 pMain :: Parser Program
652 pMain
653   = do
654       whiteSpace
655       p <- pProgram
656       eof
657       return p
658
659 main :: IO ()
660 main
661   = do { progname <- getProgName
662        ; args <- getArgs
663        ; checkArgs progname args
664        ; input <- readFile (head args)
665        ; let output = runParser pMain 0 "" input
666        ; case output of
667            Right ast -> print ast
```

```
668             Left  err -> do { putStr "Parse error at "
669                              ; print err
670                              }
671         }
672
673 checkArgs :: String -> [String] -> IO ()
674 checkArgs _ [filename]
675    = return ()
676 checkArgs progname _
677    = do { putStrLn ("Usage: " ++ progname ++ " filename\n\n")
678         ; exitWith (ExitFailure 1)
679         }
```