

Progetto Linguaggi e Compilatori 2

Parte 3

Gruppo 1

Studenti:

Foschiani Luca
Marinato Riccardo
Passalenti Andrea

1 Assunzioni

Nel seguente progetto abbiamo considerato le scelte di sintassi concreta definite dal linguaggio *Go*. A seguire vengono riassunte una serie di assunzioni.

- Variabili:
 - non è consentito dichiarare delle variabili con lo stesso identificatore nello stesso blocco, mentre è possibile se avviene su blocchi (e sottoblocchi) diversi;
 - non è possibile utilizzare delle right-expression isolate;
 - è consentita la dichiarazione di variabili globali al di fuori delle funzioni, ma non tramite la sintassi `:=`.
- Funzioni e procedure:
 - non è possibile dichiarare più di una funzione/procedura con lo stesso identificatore;
 - nel corpo delle procedure non deve presentare l'istruzione `return`, sempre richiesta invece nei corpi delle funzioni;
 - è consentito effettuare dei passaggi di parametri per valore (*val*) e per riferimento (*ref*).
- Comandi di controllo sequenza:
 - è stato usato l'operatore `!` nelle condizioni degli `if` per veicolare correttamente i salti condizionati;
 - sono stati implementati solamente i comandi per il controllo della sequenza richiesti nel testo del progetto (*condizionali semplici* e *iterazione indeterminata*);
 - le guardie booleane dei controlli di sequenza vengono gestite tramite short-cut mentre le altre espressioni booleane vengono gestite senza short-cut;
 - non è possibile definire controlli di sequenza al di fuori delle funzioni/procedure.

2 Scelte implementative

- Abbiamo generato *lexer* e *parser* tramite il tool BNFC, partendo dalla definizione di una grammatica iniziale;
- la gestione del *Type Checker* e del *Three Address Code* vengono fatte all'interno del parser;
- le funzioni *write* sono trattate come statement che prendono come argomento una right-expression, mentre le funzioni *read* vengono viste come right-expressions e hanno una lista di parametri di input vuota.

3 Grammatica

Di seguito viene riportata la grammatica di partenza.

```
comment "//";  
comment "/*" "*/";
```

```
entrypoints Start;
```

```
rules      Boolean ::= "true" | "false" ;
```

```
ExpAnd.      RExp ::= RExp "&&" RExp ;  
ExpOr.       RExp ::= RExp "||" RExp ;  
ExpNot.      RExp ::= "!" RExp ;  
ExpEq.       RExp ::= RExp "==" RExp ;  
ExpNeq.      RExp ::= RExp "!=" RExp ;  
ExpLt.       RExp ::= RExp "<" RExp ;  
ExpLtE.      RExp ::= RExp "<=" RExp ;  
ExpGt.       RExp ::= RExp ">" RExp ;  
ExpGtE.      RExp ::= RExp ">=" RExp ;  
ExpAdd.      RExp ::= RExp "+" RExp ;  
ExpSub.      RExp ::= RExp "-" RExp ;  
ExpMul.      RExp ::= RExp "*" RExp ;  
ExpDiv.      RExp ::= RExp "/" RExp ;  
ExpMod.      RExp ::= RExp "%" RExp ;  
ExpNeg.      RExp ::= "-" RExp ;
```

```

ExpRef.      RExp ::= "&" LExp ;
ExpFuncEmpty. RExp ::= Id "(" ")" ;
ExpFunc.     RExp ::= Id "(" [RExp] ")" ;
ExpVal.      RExp ::= Val ;
ExpLExp.     RExp ::= LExp ;
ExpPar.      RExp ::= "(" RExp ")";
StRead.      RExp ::= ReadT "(" ")";

Int .        Val ::= Integer ;
Float .     Val ::= Double ;
Char .      Val ::= Char ;
String .    Val ::= String ;
Bool .      Val ::= Boolean ;

rules      ReadT ::= "readInt" | "readFloat" | "readChar" | "readString" ;

ExpId.      LExp ::= Id ;
ExpArr.     LExp ::= LExp "[" RExp "]" ;
ExpDeref.   LExp ::= "*" RExp ;

separator nonempty RExp "," ;

Entry.      Start ::= "package" Id [Decl];

separator Decl "";

DeclVar.     Decl ::= "var" [Id] Type;
DeclVarInit. Decl ::= "var" [Id] "=" [RExp];
DeclVarTypeInit. Decl ::= "var" [Id] Type "=" [RExp];
DeclVarShort. ShortVarDecl ::= [Id] ":@" [RExp];

TVoid.      Type ::= "void";
TInt.       Type ::= "int";
TBool.      Type ::= "bool";
TFloat.     Type ::= "float";
TChar.      Type ::= "char";
TString.    Type ::= "string";
TArray.     Type ::= "[" Integer "]" Type;

```

```

TPointer.      Type ::= "*" Type;

token Id (letter | '_' )(letter | digit | '_' )*;

separator Param ",";

DeclFun.      Decl ::= "func" Id "(" [Param] ")" Type Block;
DeclProc.     Decl ::= "func" Id "(" [Param] ")" "void" Block;

separator nonempty Id ",";

Parameter.    Param ::= [Id] Type;
ParameterPass. Param ::= Pass [Id] Type;

PassVal.      Pass  ::= "val";
PassRef.      Pass  ::= "ref";

separator Stmt "";

BodyBlock.    Block ::= "{" [Stmt] "}";

StDecl.      Stmt ::= Decl;
StBlock.     Stmt ::= Block;
StSmpl.      Stmt ::= StmtSmpl;
StIf.        Stmt ::= "if" RExp Block;
StIfElse.    Stmt ::= "if" RExp Block "else" Block;
StWhile.     Stmt ::= "for" RExp Block;
StBreak.     Stmt ::= "break";
StContinue.  Stmt ::= "continue";
StReturn.    Stmt ::= "return" RExp;
StWrite.     Stmt ::= WriteT "(" RExp ")";

rules WriteT ::= "writeInt" | "writeFloat" | "writeChar" | "writeString";

StShortVarDecl. StmtSmpl ::= ShortVarDecl;
StExp.          StmtSmpl ::= RExp;
StAsgn.         StmtSmpl ::= LExp "=" RExp;

```

4 Parser

Gran parte del progetto è stato sviluppato all'interno del parser. In particolare tutta la parte di *type checking* e *three address code* sono state implementate al suo interno.

All'interno del parser sono stati definiti tutta una serie di attributi relativi alla *grammatica attributata*. Gli environment sono gestiti attraverso gli attributi *EnvVar/EnvVarNew* ed *EnvFun/EnvFunNew* e corrispondono a liste di tipi di dato *ElemVar* ed *ElemFun*. In particolare, gli attributi *EnvVar* ed *EnvFun* rappresentano, per ogni nodo, l'ambiente derivante dal nodo padre, mentre *EnvVarNew* ed *EnvFunNew* definiscono l'ambiente ottenuto da tutti i nodi figli. Sono stati inoltre definiti ulteriori attributi per gestire la generazione del three address code (si rimanda al codice).

Per quanto riguarda la dichiarazione delle precedenze e l'associatività degli operatori, sono tutte state gestite a livello di parser.

4.1 Type Checker

Il Type Checker viene gestito interamente all'interno del parser. In particolare per ogni espressione valutata si effettua un controllo diretto attraverso le funzioni *Ok* e *Bad* che, in caso di errore, gestiscono l'errore tramite la stampa di un messaggio a video indicante la linea dell'errore e la causa.

Un errore di lessico viene a distinguersi da uno di tipo direttamente in fase di stampa dello dell'errore, riportando "*Lexical error*" nel primo caso e "*Type error*" nel secondo. In particolare il type checker si occuperà di intercettare eventuali errori di tipo e di scoping.

Tutte le funzioni di controllo utilizzate dal Type Checker sono state definite internamente al parser.

Tra le diverse assunzioni fatte in fase di type checking, abbiamo voluto tenere conto in modo automatico di una conversione dal tipo int al tipo float.

4.2 Three Address Code

Il Three Address Code viene generato interamente nel parser, utilizzando un modulo esterno (*TAC.hs*) contenente la struttura dati per gestirne la generazione e le funzioni addette alla stampa. Inoltre ci siamo appoggiati su un ulteriore modulo (*Env.hs*) contenete tutte le funzioni per la gestione dell'environment (inserimento, cancellazione e ricerca di variabili/funzioni).

In questo modulo sono stati dichiarati i tipi di dato *ElemVar* e *ElemFun* contenenti rispettivamente tutte le informazioni relative alle variabili ed alle funzioni/procedure dichiarate nell'ambiente.

4.3 Test Case

Sono stati preparati dei test-case significativi che possono essere eseguiti in sequenza attraverso il comando *make demo*, oppure singolarmente attraverso i comandi *make demo1*, *make demo2*, *make demo3* e *make demo4*.