# תורת הקומפילציה

## תרגיל בית 2 – בנית מנתח תחבירי

mip@cs.technion.ac.il – מתרגל אחראי: מתן פלד

ההגשה בזוגות

עבור כל שאלה על התרגיל, יש לעין ראשית **בפיאצה** ובמידה שלא פורסמה אותה השאלה, ניתן להוסיף אותה ולקבל מענה, אין לשלוח מיילים בנושא התרגיל בית כדי שנוכל לענות על השאלות שלכם ביעילות.

תיקונים לתרגיל יסומנו בצהוב, חובתכם להתעדכן בהם באמצעות קובץ התרגיל.

התרגיל ייבדק בבדיקה אוטומטית. **הקפידו למלא אחר ההוראות במדויק.** 

### כללי

בתרגיל זה עליכן לממש ניתוח תחבירי לשפת FanC, הכוללת פעולות אריתמטיות, פונקציות, והמרות מובנות מ-byte (בית אחד) ל-int (4 בתים).

מנתח זה ישמש אתכם גם בתרגילים הבאים ולכן המלצתנו היא שתממשו ותגישו תרגיל זה על אף שהוא מגן.

## מנתח לקסיקלי

יש לכתוב מנתח לקסיקלי המתאים להגדרות הבאות:

אסימון
VOID
INT
BYTE
В
BOOL
AND
OR
NOT
TRUE
FALSE
RETURN
IF
ELSE
WHILE
BREAK
CONTINUE
SC
COMMA
LPAREN
RPAREN
LBRACE
RBRACE
ASSIGN
RELOP
BINOP
ID

0   [1-9][0-9]*	NUM
"([^\n\r\"\\] \\[rnt"\\])+"	STRING

ניתן לשנות את שמות האסימונים או להוסיף אסימונים נוספים במידת הצורך, כל עוד המנתח הלקסיקלי מזהה את כל התבניות לעיל.

יש להתעלם מרווחים, ירידות שורה משני הסוגים (LF ,CR) וטאבים כך שלא תתקבל עליהם שגיאה לקסיקלית. יש להתעלם מהערות שורה (הערות ++C)) המיוצגות ע"י התבנית ?[\r\n]\*[\r\n]/r\n]//

הערה: המנתח הזה שונה במקצת מהמנתח הלקסיקלי של תרגיל בית 1, אבל ניתן להשתמש במנתח שתופס אסימונים נוספים או מתמודד עם escape-ים נוספים, כמו המנתח של תרגיל בית 1, כל עוד תפסתם את כל התבניות בטבלה.

#### תחביר

יש לכתוב מנתח תחבירי שיתאים לדקדוק הבא:

- 1.  $Program \rightarrow Funcs$
- 2.  $Funcs \rightarrow \epsilon$
- 3.  $Funcs \rightarrow FuncDecl Funcs$
- 4. FuncDecl  $\rightarrow$  RetType ID LPAREN Formals RPAREN LBRACE Statements RBRACE
- 5.  $RetType \rightarrow Type$
- 6.  $RetType \rightarrow VOID$
- 7. Formals  $\rightarrow \epsilon$
- 8. Formals  $\rightarrow$  FormalsList
- 9.  $FormalsList \rightarrow FormalDecl$
- 10.  $FormalsList \rightarrow FormalDecl\ COMMA\ FormalsList$
- 11.  $FormalDecl \rightarrow Type\ ID$
- 12.  $Statements \rightarrow Statement$
- 13.  $Statements \rightarrow Statements Statement$
- 14. Statement  $\rightarrow$  LBRACE Statements RBRACE
- 15.  $Statement \rightarrow Type\ ID\ SC$
- 16.  $Statement \rightarrow Type\ ID\ ASSIGN\ Exp\ SC$
- 17. Statement  $\rightarrow$  ID ASSIGN Exp SC
- 18.  $Statement \rightarrow Call SC$
- 19.  $Statement \rightarrow RETURN SC$
- 20. Statement  $\rightarrow$  RETURN Exp SC
- 21. Statement  $\rightarrow$  IF LPAREN Exp RPAREN Statement
- 22. Statement  $\rightarrow$  IF LPAREN Exp RPAREN Statement ELSE Statement
- 23. Statement  $\rightarrow$  WHILE LPAREN Exp RPAREN Statement
- 24.  $Statement \rightarrow BREAK\ SC$
- 25. Statement  $\rightarrow$  CONTINUE SC
- 26.  $Call \rightarrow ID LPAREN ExpList RPAREN$
- 27.  $Call \rightarrow ID LPAREN RPAREN$
- 28.  $ExpList \rightarrow Exp$
- 29.  $ExpList \rightarrow Exp\ COMMA\ ExpList$
- 30.  $Type \rightarrow INT$
- 31.  $Type \rightarrow BYTE$
- 32.  $Type \rightarrow BOOL$

- 33.  $Exp \rightarrow LPAREN Exp RPAREN$
- 34.  $Exp \rightarrow Exp \ BINOP \ Exp$
- 35.  $Exp \rightarrow ID$
- 36.  $Exp \rightarrow Call$
- 37.  $Exp \rightarrow NUM$
- 38.  $Exp \rightarrow NUM B$
- 39.  $Exp \rightarrow STRING$
- 40.  $Exp \rightarrow TRUE$
- 41.  $Exp \rightarrow FALSE$
- 42.  $Exp \rightarrow NOT Exp$
- 43.  $Exp \rightarrow Exp \ AND \ Exp$
- 44.  $Exp \rightarrow Exp \ OR \ Exp$
- 45.  $Exp \rightarrow Exp \ RELOP \ Exp$
- 46.  $Exp \rightarrow LPAREN Type RPAREN Exp$

#### :הערות

- הדקדוק כפי שמוצג כאן אינו חד משמעי ב-Bison. יש להפכו לחד משמעי תוך שימור השפה. בעיה לדוגמה שיש לפתור: <a href="http://en.wikipedia.org/wiki/Dangling else">http://en.wikipedia.org/wiki/Dangling else</a>
  שיש לפתור: Dangling else ללא שינוי הדקדוק אלא באמצעות מתן עדיפות לכללים או אסוציאטיביות מתאימה לאסימונים.
- 2. יש להקפיד על מתן עדיפויות ואסוציאטיביות מתאימים לאופרטורים השונים. יש להשתמש בטבלת העדיפויות 2. http://introcs.cs.princeton.edu/java/11precedence
  - . אין צורך לבצע שינויים בדקדוק, פרט לשם הבדלה בין האופרטורים השונים.

## קלט ופלט המנתח

קובץ ההרצה של המנתח יקבל את הקלט מ-stdin.

יש להיעזר בקובץ output.hpp המצורף לתרגיל על מנת לייצר פלט הניתן לבדיקה אוטומטית.

על המנתח להדפיס את כללי הגזירה על פי סדר ביצוע פעולות הreduce. בעת ביצוע reduce לכלל יש לקרוא לפונקציה printProductionRule(ruleno)

לתרגיל מסופקים 2 דוגמאות עם קבצי קלט ופלט מתאימים. יש לבדוק שפורמט ההדפסה שלכם זהה לדוגמאות. הבדלי פורמט יגרמו לכישלון הבדיקות האוטומטיות.

## טיפול בשגיאות

בקובץ הקלט יכולות להיות שגיאות לקסיקליות ותחביריות. **על המנתח לסיים את ריצתו מיד עם זיהוי שגיאה** (כלומר בנקודה העמוקה ביותר בעץ הגזירה שבה ניתן לזהותה). ניתן להניח כי הקלט מכיל <u>שגיאה אחת לכל היותר</u>.

על מנת לדווח על שגיאות יש להשתמש בפונקציות הנתונות בקובץ output.hpp:

errorLex(lineno) שגיאה לקסיקלית errorSyn(lineno) שגיאה תחבירית

בכל השגיאות הנ"ל lineno הוא מס' השורה בה מופיעה השגיאה.

• במקרה של שגיאה הפלט של המנתח יהיה כל הכללים שנעשה להם reduce והשגיאה שהתגלתה (כפי שניתן לראות בדוגמה t2).

#### הוראות הגשה

שימו לב כי קובץ ה-Makefile מאפשר שימוש ב-STL. אין לשנות את ה-Makefile.

יש להגיש קובץ אחד בשם ID1-ID2.zip, עם מספרי ת"ז של שתי המגישות. על הקובץ להכיל:

- המכיל את כללי הניתוח הלקסיקלי scanner.lex בשם flex קובץ ∙
  - קובץ בשם parser.ypp המכיל את המנתח
- את כל הקבצים הנדרשים לבניית המנתח, כולל \*.output שסופקו כחלק מהתרגיל.

בנוסף, יש להקפיד שהקובץ לא יכיל את:

- קובץ ההרצה
- bison-ו flex קבצי הפלט של
- שסופק כחלק מהתרגיל Makefile פובץ

יש לוודא כי בביצוע unzip לא נוצרת תיקיה נפרדת. על המנתח להיבנות על השרת csComp ללא שגיאות באמצעות קובץ Makefile שסופק עם התרגיל. באתר הקורס מופיע קובץ zip המכיל קבצי בדיקה לדוגמה. יש לוודא כי פורמט הפלט זהה לפורמט הפלט של הדוגמאות הנתונות. כלומר, ביצוע הפקודות הבאות:

unzip id1-id2.zip cp path-to/Makefile . cp path-to/ hw2-tests.zip . unzip hw2-tests.zip make ./hw2 < t1.in 2>&1 > t1.res diff t1.res path-to/t1.out

ייצור את קובץ ההרצה בתיקיה הנוכחית ללא שגיאות קומפילציה, יריץ אותו, ו-diff יחזיר 0.

הגשות שלא יעמדו בדרישות לעיל יקבלו ציון 0 ללא אפשרות לבדיקה חוזרת.

בדקו היטב שההגשה שלכן עומדת בדרישות הבסיסיות הללו לפני ההגשה עצמה.

**שימו לב** כי באתר מופיע script לבדיקה עצמית לפני ההגשה בשם selfcheck. תוכלו להשתמש בו על מנת לוודא כי ההגשה שלכם תקינה.

בתרגיל זה (כמו בתרגילים אחרים בקורס) ייבדקו העתקות. אנא כתבו את הקוד שלכם בעצמכם.

בהצלחה! ☺