

## מבחן סוף סמסטר – מועד ב'

### טור 1

**מראה אחראי:** ד"ר הילה פלג

**מתרגלים:** הילה לוי, תומר כהן, אנדריי בבין, תומר ביתן

הוראות:

1. בטופס המבחן 13 עמודים, מתוכם 5 דפי נוסחאות. בדקו שכל העמודים ברשותכם.
2. משך המבחן שלוש שעות (180 דקות).
3. כל חומר עזר חיצוני אסור לשימוש.
4. בשאלות הפתוחות, ניתן לציין לגבי סעיף או שאלה "לא יודע/ת". תשובה זו תזכה ב-20% מהניקוד. תשובות שגויות לא יזכו בניקוד.
5. קראו את כל המבחן לפני שאתם מתחילים לענות על השאלות.
6. אין צורך להגיש את טופס מבחן זה בתום הבחינה.
7. את התשובות לשאלות הסגורות יש לסמן בטופס הטורי הנפרד בלבד ואת התשובות לשאלות הפתוחות יש לכתוב בריבועים הריקים לכל שאלה בטופס התשובות הנפרד.
8. ודאו כי אתם מגישים טופס תשובות וטופס טורי בלבד (את מחברת הטיוטה ניתן לשמור אצלכם).

**בהצלחה!**

## חלק א' - שאלות סגורות (45 נק')

שלבי קומפילציה (27 נק')

1,2,3

נתונה התוכנית הבאה בשפת FanC:

```
1. int magicVar = 13;  
2. int i = 10; b  
3. int k = 2360360;  
4.  
5. while (i < 300) {  
6.     k = k / i;  
7.     i = i + 1b;  
8. }  
9.  
10. printi(k);
```

בסעיפים הבאים (שאלות 1 עד 5) מוצגים שינויים (בלתי תלויים) לקוד של התוכנית. ניתן להניח כי התכנית המקורית מתקמפלת בהצלחה. עבור כל שינוי ציינו את השלב המוקדם ביותר שבו נגלה את השגיאה:

### שאלה מספר 1:

(3 נק') מחליפים את שורה 1 בשורה 1b

```
1. int magic-var = 13;
```

- א. שגיאה בייצור קוד
- ב. שגיאה בניתוח לקסיקלי
- ג. שגיאה בניתוח סמנטי
- ד. שגיאה בזמן ריצה
- ה. אין שגיאה
- ו. שגיאה בניתוח תחבירי

### שאלה מספר 2:

(3 נק') מחליפים את שורה 1 בשורה

```
1. int magic_var = 13;
```

- א. אין שגיאה
- ב. שגיאה בייצור קוד
- ג. שגיאה בניתוח תחבירי
- ד. שגיאה בזמן ריצה
- ה. שגיאה בניתוח סמנטי
- ו. שגיאה בניתוח לקסיקלי

### שאלה מספר 3:

(3 נק') מחליפים את שורה 2 בשורה

2. byte i = 10b;

- א. שגיאה בניתוח סמנטי
- ב. אין שגיאה
- ג. שגיאה בזמן ריצה
- ד. שגיאה בניתוח לקסיקלי
- ה. שגיאה בניתוח תחבירי
- ו. שגיאה בייצור קוד

### שאלה מספר 4:

(3 נק') מחליפים את שורה 3 בשורה

3. byte k = 2360360;

- א. שגיאה בניתוח תחבירי
- ב. שגיאה בניתוח לקסיקלי
- ג. שגיאה בייצור קוד
- ד. שגיאה בזמן ריצה
- ה. שגיאה בניתוח סמנטי
- ו. אין שגיאה

### שאלה מספר 5:

(3 נק') מחליפים את שורה 7 בשורה

7. i++;

- א. שגיאה בניתוח סמנטי
- ב. שגיאה בניתוח תחבירי
- ג. אין שגיאה
- ד. שגיאה בייצור קוד
- ה. שגיאה בניתוח לקסיקלי
- ו. שגיאה בזמן ריצה

### שינויים ב-FanC:

נרצה להוסיף לשפת FanC את האופרטור = : כך שהביטוי  
 $x := y$   
יכניס ל-x את הערך של y אם השמה של x ל-y היא חוקית ואז יחזיר true, ואחרת לא ישנה את ערך המשתנה ויחזיר false.  
כך שקטע הקוד הבא :

```
12. if (x := y || x:= y+1) {  
13.   print("success");  
14. }
```

- ידפיס success אם x ו-y מוגדרים כ-int, וערכו של x אחרי שורה 14 יהיה אחד יותר מערכו של y.
- ידפיס success אם x ו-y מוגדרים כ-byte, וערכו של x אחרי שורה 14 יהיה כערכו של y.
- לא ידפיס דבר אם x מוגדר כ-byte ו-y מוגדר כ-int, וערכו של x לא ישתנה.

שימו לב כי אם ההשמה הראשונה אינה חוקית, ריצת התוכנית תמשיך לנסות את ההשמה השנייה. שגיאות זמן ריצה בחישוב צד ימין של ההשמה לא משפיעות על ערך ההחזרה אלא נזרקות החוצה מן האופרטור.

### שאלה מספר 6:

(6 נק') אילו שלבים של הקומפילר נצטרך לשנות כדי לתמוך באופרטור החדש?

- ☒ א. תחבירי, סמנטי
- ☒ ב. לקסיקלי, תחבירי, סמנטי, ייצור קוד
- ☒ ג. סמנטי, ייצור קוד
- ☐ ד. לקסיקלי, תחבירי, סמנטי
- ☒ ה. לקסיקלי, ייצור קוד
- ☐ ו. לקסיקלי, תחבירי

### שאלה מספר 7:

(6 נק') כיצד ניתן למנוע מפעולות קוד ביניים לא תקינות סמנטית להתבצע בזמן חישוב אופרטור = : בזמן ריצה?

- ☒ א. אין דרך לייצר עבור האופרטור קוד שלא עלול לגרום להתנהגות בלתי-מוגדרת של התכנית
- ☒ ב. לא ייתכנו פעולות לא תקינות סמנטית בקוד ביניים, שכן אין בו טיפוסים
- ☒ ג. בדיקות הטיפוסים של שפת הביניים ימנעו מהפעולות להתבצע
- ☒ ד. ניתן להסתמך על ערך החזרה של הביטוי בזמן ייצור הקוד ולא לייצר פעולות לא תקינות סמנטית
- ☒ ה. ניתן להסתמך על ערך החזרה של הביטוי בשלב התחבירי ולהחליף את האופרטור בקבוע במחסנית של בייזון

(18 נק') אופטימיזציות

נתונות שורות הקוד הבאות שמהוות בלוק בסיסי בתוך תכנית גדולה יותר :

1. $y = 1 + 4$	<del><math>y = 5</math></del>	(8)
2. $x = 5$	<del><math>x = 5</math></del>	
3. $c = 1$	<del><math>c = 1</math></del>	
4. $k = y * c$	<del><math>k = 5</math></del>	
5. $x = g + y$	<del><math>x = g + 5</math></del>	
6. $c = c + 1$	<del><math>c = 2</math></del>	
7. $z = g + k$	<del><math>z = x</math></del>	
8. $c = z - 2$	<del><math>c = x - 2</math></del>	

ללא כל מידע נוסף, הריצו עליהן עד להתכנסות את האופטימיזציות הבאות :

- constant propagation
- copy propagation
- constant folding
- algebraic simplification
- common subexpression elimination
- useless code elimination

לאחר מכן ענו על השאלות הבאות.

**שאלה מספר 8:**

(6 נק') אם נתונה לכם לצורך שאלה זו בלבד הרצה של אנליזת חיות המוצאת שבסוף הבלוק המשתנים  $k, y$  מתים, כמה שורות יישארו בקטע הקוד אחרי ביצוע useless code elimination אחרי האופטימיזציות שכבר בוצעו?

- א. 4
- ב. 5
- ג. 2
- ד. 6
- ה. 3

**שאלה מספר 9:**

(6 נק') אחרי הרצת האופטימיזציות ללא מידע נוסף, הפקודה שהייתה במקור בשורה 7 תהיה כעת  $z = \underline{\hspace{2cm}}$  (השלימו את החסר)

א. ניתן למחוק שורה זו לגמרי

ב.  $x$

ג.  $g + k$

ד. 5

ה.  $g + 5$

ו.  $g + y$

**שאלה מספר 10:**

6 נק') אחרי הרצת האופטימיזציות ללא מידע נוסף, הפקודה שהייתה במקור בשורה 8 תהיה כעת  $C = \underline{\hspace{2cm}}$  (השלימו את החסר)

א.  $g - 3$

ב.  $x - 2$

ג.  $x - c$

ד.  $3$

ה.  $z - c$

ו. ניתן למחוק שורה זו לגמרי

**חלק ב' - שאלות פתוחות (55 נק')**

**שאלה 1: דקדוקים (25 נק')**

	S	T	#	E	R
0	s1			2	3
1	r3	r3	r3		
2	acc	acc	acc		
3	s4	s5			
4	r2	r2	r2		
5		s6			
6	r1	r1	r1		

איתם אסימונים ל-flex .  
 , ואם מילה מתאימה

נתון הדקדוק  $G$  מעל הטרמינלים  $s, t$  והמשתנים  $E, R$ , אשר נמצא  $LR(1)$ :

USF

$$E \rightarrow R t t$$
$$2 \quad R \rightarrow R, s \mid \text{[scribble]}$$
$$\exists R \rightarrow S$$
$$\{E.\text{count} = 2, R.\text{count} = E.\text{count}\}$$
$$\{ R_{i, \text{count}} = R_{\text{count}} - L_i \}$$

```

    if (R.count > 0) error();
}

```

א. (4 נק') מהי השפה המתקבלת מהדקדוק? תנו ביטוי רגולרי אחד בהתאם לכללים שהדרתם איתם אסימונים ל-lex. הביטוי צריך להתאים לכל המילים בשפה ולחן בדיק: כל מילה בשפה תתאים לביטוי הרגולרי, ואם מילה מתאימה לביטוי הרגולרי היא בהכרח בשפה.

$$S^+H$$

ב. (8 נק') מבין המחלקות LR(0), SLR, LR(1) מהי המחלקה הקטנה ביותר בה נמצא הדקדוק? תזכורת: מחלקה LR(0) קטנה יותר ממחלקה B אם מחלקה B יכולה לטפל בכל הדקדוקים בהם יכולה לטפל מחלקה A. נמקו את תשובתכם.

$$sc) (F \rightarrow R) \vdash S$$
$$sc1(R-RS) : \{s\} \cap \neq \emptyset$$
$$SC1 (K \rightarrow S) \vdash S \quad 1$$

ג. (2 נק') האם הדקדוק נמצא ב-LL(1)?

ד. (6 נק') נרצה לוודא שכמות ה-s במילה גדולה מכמות ה-t. כתבו הגדרה מונחית תחביר עבור  $G$  המשתמשת בתכונות **נורשות בלבד** שתוודא זאת. הגדירו את התכונות והפעולות הסמנטיות.

## הוראות:

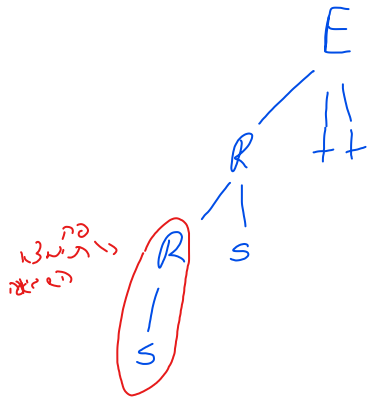
- אם המילה תקינה אין לבצע דבר, ואם המילה אינה תקינה יש לקרוא לפונקציה `error()`
- אין לשנות את הדקדוק! יש לכתוב פעולות סמנטיות רק במקומות המסומנים לכם.
- אין להוסיף לטרמינלים תכונות סמנטיות.

ציינו במפורש איזה תכונות סמנטיות הגדרתם לכל אחד מהמשתנים **ונמקו היטב!!** את משמעות התכונה.

מלאו את מימוש הבדיקה שלכם עבור כל כללי הדקדוק.

יש למלא את ההגדרות במקום המתאים לכך במחברת התשובות.

ה. (5 נק') נתחו את המילה sstt : הריצו עליה את המנתח של המחלקה הקטנה ביותר שמצאתם בסעיף ב' ואם היא תקינה תחבירית הריצו על עץ הגזירה שלה את הפתרון שלכם לסעיף ד'. מספרו את שלבי הריצה והסבירו בדיוק מתי תימצא השגיאה.



שגיאה.	השגיאה	השגיאה
acc	$(0,1)$	$s+t$
r3	$(0,1)(1,s)$	$s+t$
s4	$(0,1)(3,R)$	$s+t$
r2	$(0,1)(3,R)(u,s)$	$t$
s5	$(0,1)(3,R)$	$t$
s6	$(0,1)(3,R)(5,t)$	$t$
r1	$(0,1)(3,R)(5,t), (0,t)$	$t$
acc	$(0,1)(2,t)$	$t$

**שאלה 2: אנליזה סטטית (30 נק')**

בניסיון לשפר את אלגוריתם הקצאת הרגיסטרים של הקומפילר שלו, מקס מעוניין להכניס הפרדה בין שני סוגי משתנים: מתים ברובם (mostly dead) או מתים לגמרי (really dead). משתנה מת לגמרי הוא משתנה שאין שימוש בערך שבו באף מסלול בתכנית. משתנה מת ברובו הוא משתנה שקיים לפחות מסלול אחד בתכנית בו אין בו שימוש בערכו יותר.

א. (5 נק') לדעתו של מקס, הוא כבר למד בקורס קומפילציה אנליזה שתחשב משתנים מותם לגמרי. האם מקס צודק? כן  
אם כן, הסבירו באיזו אנליזה עליו להשתמש וכיצד עליו לשנות את הפלט שלה. אם לא, הסבירו איזו אנליזה יצטרך לשם  
כך, והגדירו את הדומיין שלה, יחס הסדר ופעולת ה-join, ופונקציית המעברים.

ב. (5 נק') הציעו למקס גם אנליזה שתמצא משתנים מתים ברובם. הגדירו את הדומיין יחס הסדר, פעולת ה-join, ופונקציית המעברים.

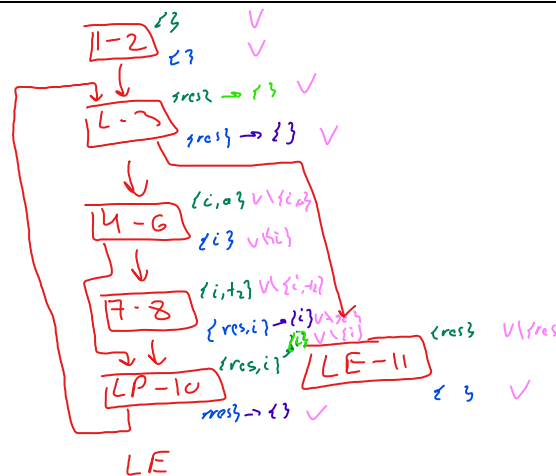
ג. (12 נק') נתונה הפונקציה הבאה.

שימו לב: הפעולה  $x^*$  קוראת מהזיכרון בכתובת השמורה ב- $x$ , כפי שראינו בהרצאה.

```

L 1. res = 0
L 2. i = 0
L loop:
L 3. if i >= 4 goto loop_end
L 4. t1 = a + i
L 5. t2 = *t1 ← gon
L 6. if t2 != 116 goto loop_post
L 7. res = t2 + 1
L 8. goto loop_post
L loop_post:
L 9. i = i + 1
L 10. goto loop
L loop_end:
L 11. print(res)

```



ציירו את ה-CFG של המונקציה והריצו את האנליזות שהגדרתם בסעיפים א' וב'. אם האנליזות שלכם דורשות חישוב נוסף על התוצאה כדי לקבל את המשתנים המתים ברובם והמתים לגמרי, כדאי לבצע אותו בצד אחרי התכנסות האנליזה.

ד. (5 נק') ציירו את גרף ה- $\text{interference}$  של הפונקציה לפי תוצאת האנליזה שלכם. להזכירכם: הצמתים הם משתנים בתכנית ואם שני משתנים לא יכולים להשתמש באותו הרגיסטר באותו הזמן מכיוון שהערכים של שניהם עדיין נחוצים בנקודה כלשהי בתכנית, תהיה ביניהם קשת בגרף.

ה. (3 נק') מקס טוען שבעת הקצאת רגיסטרים בצביעה, בחירת מועמד ל-spill **שאינו** מת ברובו או מת לגמרי יהיה עדיף על מועמד ל-spill שהוא מת ברובו. האם מקס צודק? הסבירו **בקצרה** את ההיגיון של תשובתכם על המשתנה res בשורה 3.

בהצלחה!



## נוסחאות ואלגוריתמים

כל ההגדרות מתייחסות לדקדוק  $G = (V, T, P, S)$ .

### Top Down

$$\text{first}(\alpha) = \{ t \in T \mid \alpha \Rightarrow^* t\beta \wedge \beta \in (V \cup T)^* \}$$

$$\text{follow}(A) = \{ t \in T \cup \{\$ \} \mid S\$ \Rightarrow^* \alpha A t \beta \wedge \alpha \in (V \cup T)^* \wedge \beta \in (V \cup T)^*(\epsilon | \$) \}$$

$$\text{select}(A \rightarrow \alpha) = \begin{cases} \text{first}(\alpha) \cup \text{follow}(A) & \alpha \Rightarrow^* \epsilon \\ \text{first}(\alpha) & \text{otherwise} \end{cases}$$

הגדרה: דקדוק  $G$  הוא  $LL(1)$  אם ורק אם לכל שני כללים ב- $G$  השייכים לאותו משתנה  $A$  מתקיים:  
 $\text{select}(A \rightarrow \alpha) \cap \text{select}(A \rightarrow \beta) = \emptyset$

הגדרת טבלת המעברים  $M : V \times (T \cup \{\$ \}) \rightarrow P \cup \{\text{error}\}$  עבור דקדוק  $LL(1)$ :

$$M[A, t] = \begin{cases} A \rightarrow \alpha & t \in \text{select}(A \rightarrow \alpha) \\ \text{error} & t \notin \text{select}(A \rightarrow \alpha) \text{ for all } A \rightarrow \alpha \in P \end{cases}$$

אלגוריתם מנתח  $LL(1)$ :

```
Q.push(S)
while !Q.empty() do
  X = Q.pop()
  t = next token
  if X ∈ T then
    if X = t then MATCH
    else ERROR
  else // X ∈ V
    if M[X, t] = error then ERROR
    else PREDICT(X, t)
  end if
end while
t = next token
if t = $ then ACCEPT
else ERROR
```

## Bottom Up

פריט  $LR(0)$  הוא  $(A \rightarrow \alpha \bullet \beta)$  כאשר  $A \rightarrow \alpha \beta \in P$   
 סגור ( $\text{closure}$ ) על קבוצת פריטים  $I$  מוגדר באופן אינדוקטיבי:  
 ○ בסיס:  $\text{closure}(I) = I$   
 ○ צעד: אם  $(A \rightarrow \alpha \bullet B \beta) \in \text{closure}(I)$ , אז לכל  $B \rightarrow \gamma \in P$ , גם  $(B \rightarrow \bullet \gamma) \in \text{closure}(I)$   
 פונקציית המעברים של האוטומט:

$$\delta(I, X) = \bigcup \left\{ \text{closure}(A \rightarrow \alpha X \bullet \beta) \mid (A \rightarrow \alpha \bullet X \beta) \in I \right\}$$

פריט  $LR(1)$  הוא  $(A \rightarrow \alpha \bullet \beta, t)$  כאשר  $A \rightarrow \alpha \beta \in P$ ,  $t \in T \cup \{\$ \}$   
 סגור ( $\text{closure}$ ) על קבוצת פריטים  $I$  מוגדר באופן אינדוקטיבי:  
 בסיס:  $\text{closure}(I) = I$   
 צעד: אם  $(A \rightarrow \alpha \bullet B \beta, t) \in \text{closure}(I)$ , אז לכל  $B \rightarrow \gamma \in P$  ולכל  $x \in \text{first}(\beta t)$ , גם  $(B \rightarrow \bullet \gamma, x) \in \text{closure}(I)$   
 פונקציית המעברים של האוטומט:

$$\delta(I, X) = \bigcup \left\{ \text{closure}(A \rightarrow \alpha X \bullet \beta, t) \mid (A \rightarrow \alpha \bullet X \beta, t) \in I \right\}$$

הגדרת טבלת action למנתח SLR:

$$\text{action}[i, t] = \begin{cases} \text{SHIFT}_j & \delta(I_i, t) = I_j \\ \text{REDUCE}_k & \text{rule } k \text{ is } A \rightarrow \alpha, (A \rightarrow \alpha \bullet) \in I_i \text{ and } t \in \text{follow}(A) \\ \text{ACCEPT} & (S' \rightarrow S \bullet) \in I_i \text{ and } t = \$ \\ \text{ERROR} & \text{otherwise} \end{cases}$$

הגדרת טבלת action למנתח LR(1):

$$\text{action}[i, t] = \begin{cases} \text{SHIFT}_j & \delta(I_i, t) = I_j \\ \text{REDUCE}_k & \text{rule } k \text{ is } A \rightarrow \alpha \text{ and } (A \rightarrow \alpha \bullet, t) \in I_i \\ \text{ACCEPT} & (S' \rightarrow S \bullet, \$) \in I_i \text{ and } t = \$ \\ \text{ERROR} & \text{otherwise} \end{cases}$$

הגדרת טבלת goto למנתח SLR ו-LR(1):

$$\text{goto}[i, X] = \begin{cases} j & \delta(I_i, X) = I_j \\ \text{error} & \text{otherwise} \end{cases}$$

אלגוריתם מנתח shift/reduce:

```
Q.push(0)           // where 0 is the initial state of the prefix automaton
while true do
    k = Q.top().state
    t = next token
    do action[k , t]
end while
```

## קוד ביניים

סוגי פקודות בשפת הביניים:

- |                            |                                |
|----------------------------|--------------------------------|
| <b>x := y op z</b>         | 1. משפטי השמה עם פעולה בינארית |
| <b>x := op y</b>           | 2. משפטי השמה עם פעולה אונרית  |
| <b>x := y</b>              | 3. משפטי העתקה                 |
| <b>goto L</b>              | 4. קפיצה בלתי מותנה            |
| <b>if x relop y goto L</b> | 5. קפיצה מותנה                 |
| <b>print x</b>             | 6. הדפסה                       |

## Data-Flow Analysis

ההגדרות מתייחסות ל-CFG מהצורה  $G = (V, E)$ :

הצורה הכללית של המשוואות בחישוב סריקה קדמית:

$$\begin{aligned} \text{in}(B) &= \bigsqcup_{(S,B) \in E} \text{out}(S) \\ \text{out}(B) &= f_B(\text{in}(B)) \end{aligned}$$

הצורה הכללית של המשוואות בחישוב סריקה אחורית:

$$\begin{aligned} \text{in}(B) &= \bigsqcup_{(B,D) \in E} \text{out}(D) \\ \text{out}(B) &= f_B(\text{in}(B)) \end{aligned}$$

# שפת FanC

אסימונים:

תבנית	אסימון
int	INT
byte	BYTE
b	B
bool	BOOL
and	AND
or	OR
not	NOT
true	TRUE
false	FALSE
return	RETURN
if	IF
else	ELSE
while	WHILE
break	BREAK
continue	CONTINUE
;	SC
(	LPAREN
)	RPAREN
{	LBRACE
}	RBRACE
=	ASSIGN
==   !=   <   >   <=   >=	RELOP
+   -   *   /	BINOP
[a-zA-Z][a-zA-Z0-9]*	ID
0   [1-9][0-9]*	NUM
"([^\n\r\"\\] \\[rnt\"\\])+"	STRING

1.  $Program \rightarrow Statements$
2.  $Statements \rightarrow Statement$
3.  $Statements \rightarrow Statements Statement$
4.  $Statement \rightarrow LBRACE Statements RBRACE$
5.  $Statement \rightarrow Type ID SC$
6.  $Statement \rightarrow Type ID ASSIGN Exp SC$
7.  $Statement \rightarrow ID ASSIGN Exp SC$
8.  $Statement \rightarrow Call SC$
9.  $Statement \rightarrow RETURN SC$
10.  $Statement \rightarrow IF LPAREN Exp RPAREN Statement$
11.  $Statement \rightarrow IF LPAREN Exp RPAREN Statement ELSE Statement$
12.  $Statement \rightarrow WHILE LPAREN Exp RPAREN Statement$
13.  $Statement \rightarrow BREAK SC$
14.  $Statement \rightarrow CONTINUE SC$
15.  $Call \rightarrow ID LPAREN Exp RPAREN$
16.  $Type \rightarrow INT$
17.  $Type \rightarrow BYTE$
18.  $Type \rightarrow BOOL$
19.  $Exp \rightarrow LPAREN Exp RPAREN$
20.  $Exp \rightarrow Exp BINOP Exp$
21.  $Exp \rightarrow ID$
22.  $Exp \rightarrow Call$
23.  $Exp \rightarrow NUM$
24.  $Exp \rightarrow NUM B$
25.  $Exp \rightarrow STRING$
26.  $Exp \rightarrow TRUE$
27.  $Exp \rightarrow FALSE$
28.  $Exp \rightarrow NOT Exp$
29.  $Exp \rightarrow Exp AND Exp$
30.  $Exp \rightarrow Exp OR Exp$
31.  $Exp \rightarrow Exp RELOP Exp$
32.  $Exp \rightarrow LPAREN Type RPAREN Exp$