



تمرین سری دوم

دکتر سعید پارسا

الناز رضایی

اردیبهشت ۱۴۰۱

سوال ۱:

گرامر مستقل از متن G به شرح زیر مفروض است. تعیین کنید چند زوج قاعده ی تولید در این گرامر با شرط $LL(2)$ بودن سازگار نیست؟

$$S \rightarrow ASB$$

$$S \rightarrow \lambda$$

$$A \rightarrow aA$$

$$A \rightarrow \lambda$$

$$B \rightarrow bB$$

$$B \rightarrow \lambda$$

پاسخ ۱:

می دانیم شرط $LL(k)$ بودن، این است که اشتراک مجموعه های $First$ و $Follow$ آن، کمتر از k باشد. بنابراین شرط $LL(2)$ بودن، این است که اشتراک مجموعه های $First$ و $Follow$ کمتر از ۲ باشد. طبق توضیحات فوق داریم:

- $First(S) = First(A) + First(S) + First(B) = \{a, b\}$
 $Follow(S) = Follow(B) + Follow(S) + Follow(A) = \{a, b, \lambda\}$
 $First(S) \cap Follow(S) = \{a, b\} \geq 2$
- $First(A) = \{a\}$
 $Follow(A) = \{a, \lambda\}$
 $First(A) \cap Follow(A) = \{a\} \leq 2$
- $First(B) = \{b\}$
 $Follow(B) = \{b, \lambda\}$
 $First(B) \cap Follow(B) = \{b\} \leq 2$

همانطور که مشاهده می شود، اشتراک مجموعه های $First$ و $Follow$ برای A و B کوچک تر از ۲ می باشد، بنابراین $LL(2)$ می باشند اما مجموعه S شرط $LL(2)$ بودن را نقض می کند.

سوال ۲:

گرامر $LL(1)$ زیر را در نظر بگیرید . با استفاده از توابع $first$ ، $fallow$ ، جدول تجزیه این گرامر را رسم کنید.

$$E \rightarrow TE'$$

$$E' \rightarrow +TE' \mid \mathcal{L}$$

$$T \rightarrow FT'$$

$$T' \rightarrow *FT' \mid \mathcal{L}$$

$$F \rightarrow id \mid (E)$$

پاسخ ۲:

مجموعه‌های $first$:

$$First(E) = First(T) = First(F) = \{id, (\}$$

$$First(E') = \{+\}$$

$$First(T) = First(F) = \{id, (\}$$

$$First(T') = \{*\}$$

$$First(F) = \{id, (\}$$

مجموعه‌های $follow$:

$$Follow(E) = \{), \$\}$$

$$Follow(E') = \{ \$,) \}$$

$$Follow(T) = First(E') + \{ \$,) \} = \{ +, \$,) \}$$

$$Follow(T') = \{ +,), \$ \}$$

$$Follow(F) = First(T') + \{ +,), \$ \} = \{ *, +,), \$ \}$$

جدول تجزیہ :

	+	*	id	(\$)
E	-	-	TE'	TE'	-	-
E'	+TE'	-	-	-	λ	λ
T	-	-	FT'	FT'	-	-
T'	λ	*FT'	-	-	λ	λ
F	-	-	id	(E)	-	-

سوال ۳:

یک فایل گرامر antlr طراحی کنید که تمام رشته‌هایی که به فرم 'hello'ID هستند را تشخیص دهد. درخت تجزیه را نیز برای چند رشته مختلف به دست آورید.

پاسخ ۳:

فایل انتلر مربوط به این سوال، به فایل اصلی این تمرین ضمیمه شده است.

عکس کدهای انتلر این سوال:

```
1 grammar helloID;
2
3 compilationUnit:
4     'hello ' IDENTIFIER EOF
5 ;
6
7 IDENTIFIER:
8     Letter LetterOrDigit*
9 ;
10
11 fragment Letter
12     : [a-zA-Z$_] // these are the "java letters" below 0x7F
13     | ~[\u0000-\u007F\u0080-\u0BFF] // covers all characters above 0x7F which are not a surrogate
14     | [\u0800-\u0BFF] [\uDC00-\uDFFF] // covers UTF-16 surrogate pairs encodings for U+10000 to U+10FFFF
15 ;
16
17 fragment LetterOrDigit
18     : Letter
19     | [0-9]
20 ;
```

درخت‌های تجزیه برای چند رشته مختلف:



