

به نام خدا



نظریه زبان ها و ماشین ها

پروژه اختیاری

استاد درس : استاد احمد رضا منتظرالقائم

دستیاران آموزشی:

فائزه صالحی

مهرزاد انصاری پور

امیرحسین جعفری

کیارش گیلانیان

میعاد کیمیاگری

امیرطاها نجف

مهسا قندهاری

سنا باقری

هدف پروژه

در این پروژه می‌خواهیم برنامه‌ای بنویسیم که یک گرامر مستقل از متن (CFG) را به فرم نرمال چامسکی تبدیل کند. این فرم استاندارد برای استفاده در الگوریتم‌های تجزیه مثل CYK ضروری است و از قوانین ساده‌تری تشکیل شده که پردازش ماشینی را آسان‌تر می‌کند.

شرح پروژه

برنامه ابتدا یک گرامر را از فایل ورودی دریافت می‌کند؛ شامل مجموعه‌ای از ترمینال‌ها، متغیرها (غیرترمینال‌ها) و قوانین تولید. سپس طی چند مرحله گرامر را بازنویسی می‌کند تا به فرم نرمال چامسکی برسد.

ورودی برنامه

- یک فایل متنی شامل:
 ۱. خط اول: تعداد قوانین گرامر
 ۲. خط دوم: فهرست نمادهای پایانی (ترمینال‌ها) با فاصله
 ۳. خط سوم: فهرست متغیرها (غیرترمینال‌ها) با فاصله (نماد شروع، اولین عنصر)
 ۴. خطوط بعد: هر قانون در قالب $A \rightarrow \alpha \mid \beta \mid \dots$ در یک خط مجزا

- مراحل اصلی تبدیل به CNF به این شکل هستند:

۱. حذف قوانین میرا (ε تولیدهایی که به رشته خالی ختم می‌شوند)
۲. حذف قواعد یک‌ه
۳. حذف قوانین بی‌فایده
۴. سپس برای هر نماد پایانی a متغیر جدید T_a را تعریف می‌کنیم.
۵. تبدیل قوانین طولانی‌تر از ۲ نماد به زنجیره‌ای از تولیدهای دودویی.

در نهایت، گرامر به فرم CNF را در فایل خروجی با همان فرمت ورودی ذخیره کنید.

نمونه ورودی و خروجی

ورودی:

$S \rightarrow ASB$
 $A \rightarrow aAS \mid a \mid \varepsilon$
 $B \rightarrow SbS \mid A \mid bb$

که در فایل ورودی به ش کل زیر می‌آید:

۳
a b
S A B
S -> ASB
A -> aAS | a | @
B -> SbS | A | bb

خروجی:

$S0 \rightarrow AS \mid PB \mid SB$
 $S \rightarrow AS \mid PB \mid SB$
 $A \rightarrow RS \mid XS \mid a$
 $B \rightarrow TS \mid YY \mid RS \mid XS \mid a$
 $X \rightarrow a$
 $Y \rightarrow b$
 $P \rightarrow AS$
 $R \rightarrow XA$
 $T \rightarrow SY$

که در فایل خروجی به فرمت زیر نوشته می‌شود:

۹

a b

S0 S A B X Y P R T

$S0 \rightarrow AS \mid PB \mid SB$

$S \rightarrow AS \mid PB \mid SB$

$A \rightarrow RS \mid XS \mid a$

$B \rightarrow TS \mid YY \mid RS \mid XS \mid a$

$X \rightarrow a$

$Y \rightarrow b$

$P \rightarrow AS$

$R \rightarrow XA$

$T \rightarrow SY$

نکات تکمیلی

- توجه داشته باشید که در فایل ورودی، علامت @ به معنی ϵ است.
- از هر زیان برنامه نویسی دلخواه می‌توانید استفاده کنید.
- در فایل خروجی، قوانین را به ترتیبی چاپ کنید که متغیرشان در خط ۳ آمده است.