بسم الله الرحمن الرحیم

نظریه زبان‌ها و ماشین‌ها

جلسه ۱۵

مجتبی خلیلی
دانشکده برق و کامپیوتر
دانشگاه صنعتی اصفهان

تجزیه

Given a string w of terminals, we want to know whether or not w is in $L(G)$. If so, we may want to find a derivation of w. An algorithm that can tell us whether w is in $L(G)$ is a membership algorithm. The term **parsing** describes finding a sequence of productions by which a $w \in L(G)$ is derived.

تجزیه

$$S \to 0S1 \mid 1S0S \mid T$$

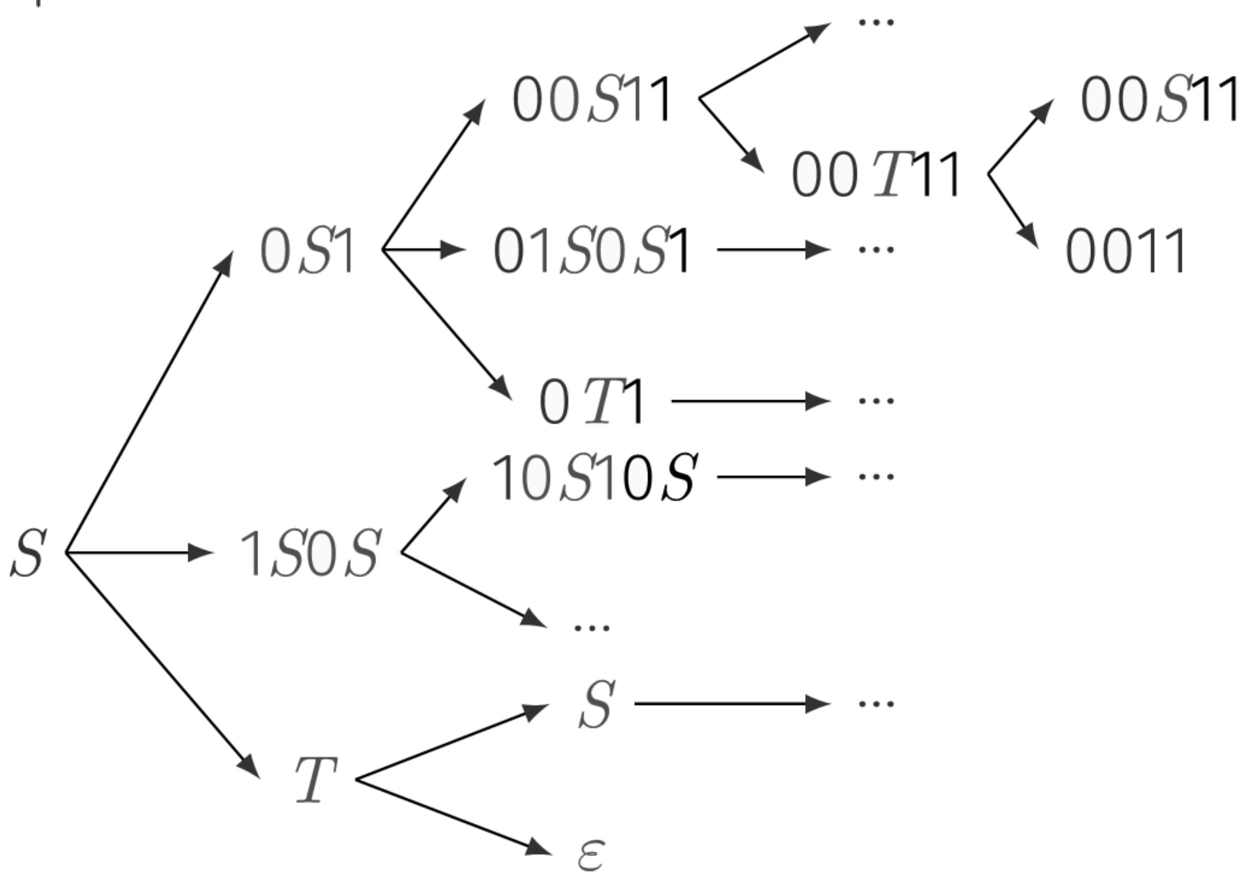
$$T \to S \mid \varepsilon$$

$$0011 \in L?$$

تجزیه

$$S \rightarrow 0S1 \mid 1S0S \mid T$$

$$T \rightarrow S \mid \varepsilon$$



تجزیه

Exhaustive search parsing has serious flaws. The most obvious one is its tediousness; it is not to be used where efficient parsing is required. But even when efficiency is a secondary issue, there is a more pertinent objection. While the method always parses a $w \in L(G)$, it is possible that it never terminates for strings not in $L(G)$.

تجزیه

- ایده: چنانچه طول رشته اشتقاق شده بیشتر از ورودی بود، ادامه نده.
- با اینحال موانعی باقی میماند:

تجزیه

$$
\begin{aligned}
S &\rightarrow AS \mid B \\
A &\rightarrow B \mid \epsilon \\
B &\rightarrow A \mid b
\end{aligned}
$$

- وجود حلقه (unit rules):

$$S \Rightarrow B \Rightarrow b$$

$$S \Rightarrow B \Rightarrow A \Rightarrow B \Rightarrow b$$

$$S \Rightarrow B \Rightarrow A \Rightarrow B \Rightarrow A \Rightarrow B \Rightarrow b$$

- ناپدید شدن متغیرها (ϵ −rules):

$$S \Rightarrow AS \stackrel{*}{\Rightarrow} AAAS \Rightarrow AAAB \stackrel{*}{\Rightarrow} B \Rightarrow b$$

$$S \Rightarrow AS \stackrel{*}{\Rightarrow} AAAAAS \Rightarrow AAAAAB \stackrel{*}{\Rightarrow} B \Rightarrow b$$

حذف برخی جملات گرامرها

Before we can study context-free languages in greater depth, we must attend to some technical matters. The definition of a context-free grammar imposes no restriction whatsoever on the right side of a production. However, complete freedom is not necessary and, in fact, is a detriment in some arguments. In Theorem 5.2, we see the convenience of certain restrictions on grammatical forms; eliminating rules of the form $A \to \lambda$ and $A \to B$ make the arguments easier. In many instances, it is desirable to place even more stringent restrictions on the grammar.

ϵ −productions حذف

DEFINITION 6.2

Any production of a context-free grammar of the form

$$A \rightarrow \lambda$$

is called a λ-**production**. Any variable A for which the derivation

$$A \stackrel{*}{\Rightarrow} \lambda \tag{6.3}$$

is possible is called **nullable**.

ϵ −productions حذف

IUT-ECE

EXAMPLE 6.4

Consider the grammar

$$S \rightarrow aS_1b,$$
$$S_1 \rightarrow aS_1b|\lambda,$$

with start variable S. This grammar generates the λ-free language $\{a^nb^n : n \geq 1\}$. The λ-production $S_1 \rightarrow \lambda$ can be removed after adding new productions obtained by substituting λ for S_1 where it occurs on the right. Doing this we get the grammar

$$S \rightarrow aS_1b|ab,$$
$$S_1 \rightarrow aS_1b|ab.$$

We can easily show that this new grammar generates the same language as the original one.

In more general situations, substitutions for λ-productions can be made in a similar, although more complicated, manner.

حذف $\epsilon -$productions

- سپس در نظر گرفتن همه nullable ها
- بازنویسی گرامر بر اساس آنها
- Start میتواند به ϵ برود.

ϵ −productions حذف

IUT-ECE

EXAMPLE 6.5

Find a context-free grammar without λ-productions equivalent to the grammar defined by

$$
\begin{aligned}
S &\rightarrow ABaC, \\
A &\rightarrow BC, \\
B &\rightarrow b|\lambda, \\
C &\rightarrow D|\lambda, \\
D &\rightarrow d.
\end{aligned}
$$

ϵ −productions حذف

From the first step of the construction in Theorem 6.3, we find that the nullable variables are A, B, C. Then, following the second step of the construction, we get

$$\begin{aligned} S &\to ABaC\,|BaC|\,AaC\,|ABa|\,aC\,|Aa|\,Ba|a, \\ A &\to B\,|C|\,BC, \\ B &\to b, \\ C &\to D, \\ D &\to d. \end{aligned}$$

مثال

- حذف ϵ −productions

$$
\begin{aligned}
S &\to 0ABC \mid 1B \mid BB \\
A &\to ABB0 \mid C \\
B &\to 0B \mid 1 \\
C &\to CC \mid \epsilon \\
D &\to 1D \mid AA
\end{aligned}
$$

$$
\begin{aligned}
S &\to 0ABC \mid 0BC \mid 0AB \mid 0B \mid 1B \mid BB \\
A &\to ABB0 \mid BB0 \mid C \\
B &\to 0B \mid 1 \\
C &\to CC \mid C \\
D &\to 1D \mid 1 \mid AA \mid A
\end{aligned}
$$

unit production حذف

DEFINITION 6.3

Any production of a context-free grammar of the form

$$A \to B,$$

where $A, B \in V$, is called a **unit-production**.

unit production حذف

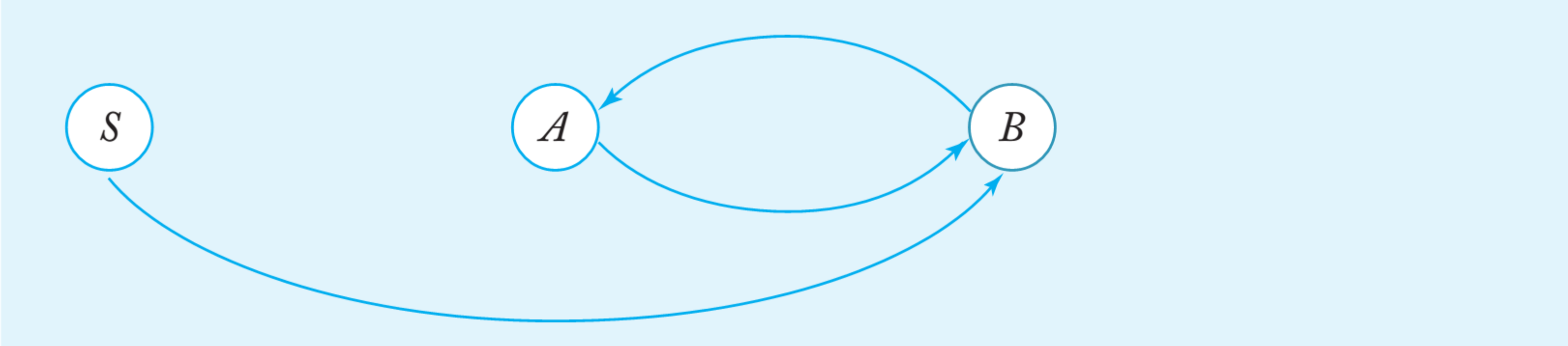
IUT-ECE

EXAMPLE 6.6

Remove all unit-productions from

$$S \rightarrow Aa|B,$$
$$B \rightarrow A|bb,$$
$$A \rightarrow a\,|bc|\,B.$$

The dependency graph for the unit-productions:



حذف unit production

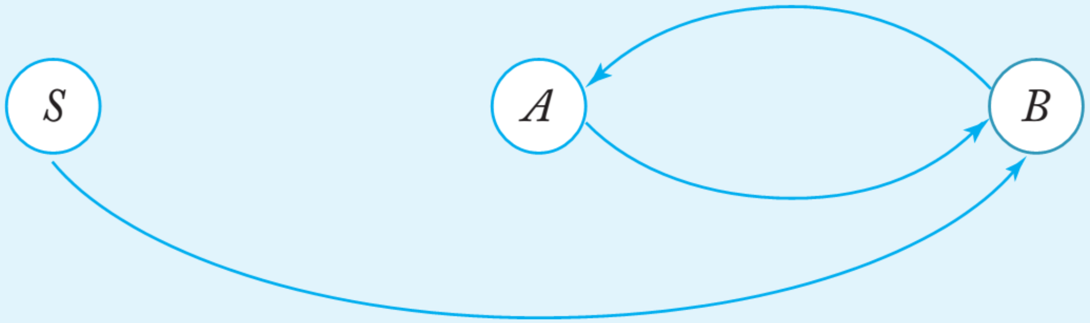
- سپس در نظر گرفتن گراف
- بازنویسی گرامر بر اساس آنها

حذف unit production

$$
\begin{aligned}
S &\to Aa \mid B,\\
B &\to A \mid bb,\\
A &\to a \mid bc \mid B.
\end{aligned}
$$

$$
\begin{aligned}
S &\to Aa,\\
A &\to a \mid bc,\\
B &\to bb,
\end{aligned}
$$

$$
\begin{aligned}
S &\to a \mid bc \mid bb \mid Aa,\\
A &\to a \mid bb \mid bc,\\
B &\to a \mid bb \mid bc.
\end{aligned}
$$



متغیرهای بی فایده

IUT-ECE

$$S \to a\,|bc|\,bb|Aa,$$
$$A \to a\,|bb|\,bc,$$
$$B \to a\,|bb|\,bc.$$

$$S \to a\,|bc|\,bb|Aa,$$
$$A \to a\,|bb|\,bc,$$

DEFINITION 6.1

Let $G = (V, T, S, P)$ be a context-free grammar. A variable $A \in V$ is said to be **useful** if and only if there is at least one $w \in L(G)$ such that

$$S \stackrel{*}{\Rightarrow} xAy \stackrel{*}{\Rightarrow} w, \tag{6.2}$$

with x, y in $(V \cup T)^*$. In words, a variable is useful if and only if it occurs in at least one derivation. A variable that is not useful is called **useless**. A production is useless if it involves any useless variable.

حذف unit production (مثال)

داریم:

$$
\begin{aligned}
S &\rightarrow 0ABC \mid 0BC \mid 0AB \mid 0B \mid 1B \mid BB \\
A &\rightarrow ABB0 \mid BB0 \mid C \\
B &\rightarrow 0B \mid 1 \\
C &\rightarrow CC \mid C \\
D &\rightarrow 1D \mid 1 \mid AA \mid A
\end{aligned}
$$

$$
\begin{aligned}
S &\rightarrow 0ABC \mid 0BC \mid 0AB \mid 0B \mid 1B \mid BB \\
A &\rightarrow ABB0 \mid BB0 \mid CC \\
B &\rightarrow 0B \mid 1 \\
C &\rightarrow CC \\
D &\rightarrow 1D \mid 1 \mid AA \mid ABB0 \mid BB0 \mid CC
\end{aligned}
$$

حذف متغیرهای بی فایده (مثال)

IUT-ECE

$$S \to 0ABC \mid 0BC \mid 0AB \mid 0B \mid 1B \mid BB$$
$$A \to ABB0 \mid BB0 \mid CC$$
$$B \to 0B \mid 1$$
$$C \to CC$$
$$D \to 1D \mid 1 \mid AA \mid ABB0 \mid BB0 \mid CC$$

داریم:

$$S \to 0AB \mid 0B \mid 1B \mid BB$$
$$A \to ABB0 \mid BB0$$
$$B \to 0B \mid 1$$