

已知文法 $G[S]$ ：

$$S \rightarrow MH \mid a$$

$$H \rightarrow LSo \mid \varepsilon$$

$$K \rightarrow dML \mid \varepsilon$$

$$L \rightarrow eHf$$

$$M \rightarrow K \mid bLM$$

1. 求每一非终结符的 FIRST 集和 FOLLOW 集。
2. 填写该文法的 LL(1) 分析表。
3. 该文法是否为 LL(1) 文法？请解释你的回答。

对于 FIRST、FOLLOW、NULLABLE 集，定义表示如下：

1. FIRST 集：

$$\text{FIRST}(A) = \{a \mid A \xRightarrow{*} aB\}$$

其中， $a \in V_T$ （终结符）， $B \in V^*$ （任意符号串）， $\xRightarrow{*}$ 表示「经由任意次推导」。

2. FOLLOW 集：

$$\text{FOLLOW}(A) = \{a \mid S \xRightarrow{*} \cdots Aa \cdots\}$$

其中 S 为开始符号， $a \in V_T$ （终结符）。

3. NULLABLE 集：假设 X 为非终结符。 $X \in \text{NULLABLE}$ ，当且仅当 $X \rightarrow \varepsilon$ 或者 $X \rightarrow \beta_1\beta_2 \cdots \beta_n$ （其中每个 β_i 均属于 NULLABLE）。

$\text{FIRST}(A)$ 就是非终结符 A 经任意次推导后，可能得到的开头终结符的集合。 $\text{FOLLOW}(A)$ 就是开始符 S 经任意次推导后， A 可能紧接着的终结符的集合。

1 求 Nullable 集、First 集

容易看出这里面有可能推出空单词 ε 的非终结符，需要先把可推空单词的非终结符之集合： **Nullable 集** 求出来。

第一遍迭代容易看出， $H, K \in \text{Nullable}$ 。第二遍迭代，因为 $K \in \text{Nullable}$ ，所以由规则 5， $M \in \text{Nullable}$ 。第三遍迭代，因为 $M, H \in \text{Nullable}$ ，所以由规则 1， $S \in \text{Nullable}$ 。因此：

$$\text{Nullable} = \{S, H, K, M\} \quad (1)$$

接下来是 **First 集**，算法如下：

Algorithm 1: First 集的计算

```
1 初始化：对所有非终结符  $N$ ， $\text{FIRST}(N) \leftarrow \emptyset$  /* 「 $\leftarrow$ 」表示赋值 */
2 while some sets changed do
3     for production  $p: N \rightarrow \beta_1\beta_2\cdots\beta_n$  do
4         for  $\beta_i \leftarrow \beta_1$  to  $\beta_n$  do /* 顺序遍历产生式的每个符号 */
5             if  $\beta_i = a$  then
6                  $\text{FIRST}(N) \leftarrow \text{FIRST}(N) \cup \{a\}$ 
7                 break
8             if  $\beta_i = M$  then
9                  $\text{FIRST}(N) \leftarrow \text{FIRST}(N) \cup \text{FIRST}(M)$ 
10                if  $M \notin \text{Nullable}$  then
11                    break
```

由 **Nullable 集**信息，得到：

- $\text{FIRST}(S) = \text{FIRST}(M) \cup \text{FIRST}(H) \cup \{a\}$
- $\text{FIRST}(H) = \text{FIRST}(L)$
- $\text{FIRST}(K) = \{d\}$
- $\text{FIRST}(L) = \{e\}$
- $\text{FIRST}(M) = \text{FIRST}(K) \cup \{b\}$

多轮迭代后得到各个非终结符的 **First 集**如下（第 4 轮结果与第 3 轮完全一致，终止）：

非终结符	S	H	K	L	M
FIRST 集 · 第 1 轮	$\{a\}$	\emptyset	$\{d\}$	$\{e\}$	$\{b, d\}$
FIRST 集 · 第 2 轮	$\{a, b, d\}$	$\{e\}$	$\{d\}$	$\{e\}$	$\{b, d\}$
FIRST 集 · 第 3 轮	$\{a, b, d, e\}$	$\{e\}$	$\{d\}$	$\{e\}$	$\{b, d\}$
FIRST 集 · 第 4 轮	$\{a, b, d, e\}$	$\{e\}$	$\{d\}$	$\{e\}$	$\{b, d\}$

表 1: FIRST 集结果

2 求 Follow 集

求 FOLLOW 集的算法如下：

Algorithm 2: FOLLOW 集的计算

```

1 初始化：对所有非终结符  $N$ ， $\text{FOLLOW}(N) \leftarrow \emptyset$ 
2 while some sets changed do
3   for production  $p: N \rightarrow \beta_1\beta_2\cdots\beta_n$  do
4      $\text{temp} \leftarrow \text{FOLLOW}(N)$ 
5     for  $\beta_i \leftarrow \beta_n$  to  $\beta_1$  do /* 逆序遍历产生式的每个符号 */
6       if  $\beta_i = a$  then
7          $\text{temp} \leftarrow \{a\}$ 
8       if  $\beta_i = M$  then
9          $\text{FOLLOW}(M) \leftarrow \text{FOLLOW}(M) \cup \text{temp}$ 
10        if  $M \notin \text{NULLABLE}$  then
11           $\text{temp} \leftarrow \text{FIRST}(M)$ 
12        else
13           $\text{temp} \leftarrow \text{FIRST}(M) \cup \text{temp}$ 

```

算法执行流程：初始时，每个非终结符的 $\text{FOLLOW} = \emptyset$ 。在整个算法中，每个非终结符的 FOLLOW 将会越来越充实，不会重置；而 temp 可能会重置。

以下表 2、表 3 是计算 FOLLOW 集的详细过程。计算到第 3 轮时，已经没有任何 FOLLOW 集发生变化。阅读表格时，请依次从上往下阅读，注意盯住每一个符号，不要看串行了。

产生式	β_i	更新 FOLLOW	更新 $temp$ 为
$S \rightarrow MH$	-	-	$FOLLOW(S) = \emptyset$
	H	$FOLLOW(H) \leftarrow FOLLOW(H) \cup \emptyset = \emptyset$	$temp \cup FIRST(H) = \{e\}$
	M	$FOLLOW(M) \leftarrow FOLLOW(M) \cup \{e\} = \{e\}$	$temp \cup FIRST(M) = \{b, d, e\}$
$S \rightarrow a$	a	-	$\{a\}$
$H \rightarrow LSo$	-	-	$FOLLOW(H) = \emptyset$
	o	-	$\{o\}$
	S	$FOLLOW(S) \leftarrow FOLLOW(S) \cup \{o\} = \{o\}$	$temp \cup FIRST(S) = \{a, b, d, e, o\}$
	L	$FOLLOW(L) \leftarrow FOLLOW(L) \cup \{a, b, d, e, o\} = \{a, b, d, e, o\}$	$FIRST(L) = \{e\}$
$H \rightarrow \varepsilon$			
$K \rightarrow dML$	-	-	$FOLLOW(K) = \emptyset$
	L	$FOLLOW(L) \leftarrow FOLLOW(L) \cup \emptyset = \{a, b, d, e, o\}$	$FIRST(L) = \{e\}$
	M	$FOLLOW(M) \leftarrow FOLLOW(M) \cup \{e\} = \{e\}$	$temp \cup FIRST(M) = \{b, d, e\}$
	d	-	$\{d\}$
$K \rightarrow \varepsilon$			
$L \rightarrow eHf$	-	-	$FOLLOW(L) = \{a, b, d, e, o\}$
	f	-	$\{f\}$
	H	$FOLLOW(H) \leftarrow FOLLOW(H) \cup \{f\} = \{f\}$	$temp \cup FIRST(H) = \{e, f\}$
	e	-	$\{e\}$
$M \rightarrow K$	-	-	$FOLLOW(M) = \{e\}$
	K	$FOLLOW(K) \leftarrow FOLLOW(K) \cup FOLLOW(M) = \{e\}$	$temp \cup FIRST(K) = \{d, e\}$
$M \rightarrow bLM$	-	-	$FOLLOW(M) = \{e\}$
	M	$FOLLOW(M) \leftarrow FOLLOW(M) \cup \{e\} = \{e\}$	$temp \cup FIRST(M) = \{b, d, e\}$
	L	$FOLLOW(L) \leftarrow FOLLOW(L) \cup \{b, d, e\} = \{a, b, d, e, o\}$	$FIRST(L) = \{e\}$
	b	-	$\{b\}$

表 2: FOLLOW 集的第 1 轮迭代

产生式	β_i	更新 FOLLOW	更新 temp 为
$S \rightarrow MH$	-	-	$\text{FOLLOW}(S) = \{o\}$
	H	$\text{FOLLOW}(H) \leftarrow \text{FOLLOW}(H) \cup \{o\} = \{f, o\}$	$\text{temp} \cup \text{FIRST}(H) = \{e, o\}$
	M	$\text{FOLLOW}(M) \leftarrow \text{FOLLOW}(M) \cup \{e, o\} = \{e, o\}$	$\text{temp} \cup \text{FIRST}(M) = \{b, d, e, o\}$
$S \rightarrow a$	a	-	$\{a\}$
$H \rightarrow LSo$	-	-	$\text{FOLLOW}(H) = \{f, o\}$
	o	-	$\{o\}$
	S	$\text{FOLLOW}(S) \leftarrow \text{FOLLOW}(S) \cup \{o\} = \{o\}$	$\text{temp} \cup \text{FIRST}(S) = \{a, b, d, e, o\}$
	L	$\text{FOLLOW}(L) \leftarrow \text{FOLLOW}(L) \cup \{a, b, d, e, o\} = \{a, b, d, e, o\}$	$\text{FIRST}(L) = \{e\}$
$H \rightarrow \varepsilon$			
$K \rightarrow dML$	-	-	$\text{FOLLOW}(K) = \{e\}$
	L	$\text{FOLLOW}(L) \leftarrow \text{FOLLOW}(L) \cup \{e\} = \{a, b, d, e, o\}$	$\text{FIRST}(L) = \{e\}$
	M	$\text{FOLLOW}(M) \leftarrow \text{FOLLOW}(M) \cup \{e\} = \{e, o\}$	$\text{temp} \cup \text{FIRST}(M) = \{b, d, e\}$
	d	-	$\{d\}$
$K \rightarrow \varepsilon$			
$L \rightarrow eHf$	-	-	$\text{FOLLOW}(L) = \{a, b, d, e, o\}$
	f	-	$\{f\}$
	H	$\text{FOLLOW}(H) \leftarrow \text{FOLLOW}(H) \cup \{f\} = \{f, o\}$	$\text{temp} \cup \text{FIRST}(H) = \{e, f\}$
	e	-	$\{e\}$
$M \rightarrow K$	-	-	$\text{FOLLOW}(M) = \{e, o\}$
	K	$\text{FOLLOW}(K) \leftarrow \text{FOLLOW}(K) \cup \text{FOLLOW}(M) = \{e, o\}$	$\text{temp} \cup \text{FIRST}(K) = \{d, e, o\}$
$M \rightarrow bLM$	-	-	$\text{FOLLOW}(M) = \{e, o\}$
	M	$\text{FOLLOW}(M) \leftarrow \text{FOLLOW}(M) \cup \{e, o\} = \{e, o\}$	$\text{temp} \cup \text{FIRST}(M) = \{b, d, e, o\}$
	L	$\text{FOLLOW}(L) \leftarrow \text{FOLLOW}(L) \cup \{b, d, e, o\} = \{a, b, d, e, o\}$	$\text{FIRST}(L) = \{e\}$
	b	-	$\{b\}$

表 3: FOLLOW 集的第 2 轮迭代 (红色文字代表相比于第 1 轮的 FOLLOW 集的变化)

最终，计算得到的 FOLLOW 集表如下表 4 所示。能坚持到这里，很不容易，不妨给自己点个赞吧！

非终结符	<i>S</i>	<i>H</i>	<i>K</i>	<i>L</i>	<i>M</i>
FOLLOW 集 · 第 1 轮	{ <i>o</i> }	{ <i>f</i> }	{ <i>e</i> }	{ <i>a, b, d, e, o</i> }	{ <i>e</i> }
FOLLOW 集 · 第 2 轮	{ <i>o</i> }	{ <i>f, o</i> }	{ <i>e, o</i> }	{ <i>a, b, d, e, o</i> }	{ <i>e, o</i> }
FOLLOW 集 · 第 3 轮	{ <i>o</i> }	{ <i>f, o</i> }	{ <i>e, o</i> }	{ <i>a, b, d, e, o</i> }	{ <i>e, o</i> }

表 4: FOLLOW 集结果

3 求 First_s 集，下结论

FIRST_s 集（有时也叫 SELECT 集），表示一条产生式可以推出来的第一个终结符的集合。

Algorithm 3: FIRST_S 集的计算

输入: 产生式 $p : N \rightarrow \beta_1\beta_2\cdots\beta_n$

输出: FIRST_S(p)

```
1 for  $\beta_i \leftarrow \beta_1$  to  $\beta_n$  do /* 顺序遍历产生式的每个符号 */
2   if  $\beta_i = a$  then
3     FIRST_S( $p$ )  $\leftarrow$  FIRST_S( $p$ )  $\cup$  { $a$ }
4   return
5   if  $\beta_i = M$  then
6     FIRST_S( $p$ )  $\leftarrow$  FIRST_S( $p$ )  $\cup$  FIRST( $M$ )
7     if  $M \notin \text{NULLABLE}$  then
8       return
9 FIRST_S( $p$ )  $\leftarrow$  FIRST_S( $p$ )  $\cup$  FOLLOW( $N$ )
10 return
```

通过先前的 FIRST 和 FOLLOW 集（集中列出于表 5），可以得到每个产生式的 FIRST_S 集，如表 6 所示。

于是，我们就可以填写 LL(1) 分析表了，如表 7 所示。

	FIRST	FOLLOW
S	$\{a, b, d, e\}$	$\{o\}$
H	$\{e\}$	$\{f, o\}$
K	$\{d\}$	$\{e, o\}$
L	$\{e\}$	$\{a, b, d, e, o\}$
M	$\{b, d\}$	$\{e, o\}$

表 5: FIRST 和 FOLLOW 集结果

	产生式	展开式	结果
1	$\text{FIRST_S}(S \rightarrow MH)$	$\text{FIRST}(M) \cup \text{FIRST}(H) \cup \text{FOLLOW}(S)$	$\{b, d, e, o\}$
2	$\text{FIRST_S}(S \rightarrow a)$	-	$\{a\}$
3	$\text{FIRST_S}(H \rightarrow LSo)$	$\text{FIRST}(L)$	$\{e\}$
4	$\text{FIRST_S}(H \rightarrow \varepsilon)$	$\text{FOLLOW}(H)$	$\{f, o\}$
5	$\text{FIRST_S}(K \rightarrow dML)$	-	$\{d\}$
6	$\text{FIRST_S}(K \rightarrow \varepsilon)$	$\text{FOLLOW}(K)$	$\{e, o\}$
7	$\text{FIRST_S}(L \rightarrow eHf)$	-	$\{e\}$
8	$\text{FIRST_S}(M \rightarrow K)$	$\text{FIRST}(K) \cup \text{FOLLOW}(M)$	$\{d, e, o\}$
9	$\text{FIRST_S}(M \rightarrow bLM)$	-	$\{b\}$

表 6: FIRST_S 集结果

	a	b	d	e	f	o
S	$\rightarrow a$	$\rightarrow MH$	$\rightarrow MH$	$\rightarrow MH$	-	$\rightarrow MH$
H	-	-	-	$\rightarrow LSo$	$\rightarrow \varepsilon$	$\rightarrow \varepsilon$
K	-	-	$\rightarrow dML$	$\rightarrow \varepsilon$	-	$\rightarrow \varepsilon$
L	-	-	-	$\rightarrow eHf$	-	-
M	-	$\rightarrow bLM$	$\rightarrow K$	$\rightarrow K$	-	$\rightarrow K$

表 7: LL(1) 分析表结果

最后判别是不是 LL(1) 文法。

一个上下文无关文法是 LL(1) 文法的充分必要条件是，对每个非终结符 A 的两条产生式 $A \rightarrow \alpha$ 、 $A \rightarrow \beta$ （其中 α, β 不同时 $\xrightarrow{*} \varepsilon$ ），均有

$$\text{FIRST}_S(A \rightarrow \alpha) \cap \text{FIRST}_S(A \rightarrow \beta) = \emptyset$$

反映在分析表上，就是看是否存在同一表项中至少有 2 个产生式（即构成冲突）。如果没有冲突，那就是 LL(1) 文法；如果有冲突，就要对文法进行等价变换，比如消除左递归和提取左公因子。

通过表 7 可知，没有冲突，所以是 LL(1) 文法。

【结论】 1. 如表 5 所示；2. 如表 7 所示；3. 是 LL(1) 文法。

【点评】 这是一道编译原理的大题，考察了 LL(1) 文法的分析方法，即依次求解 NULLABLE—FIRST—FOLLOW—FIRST_S—LL(1) 分析表。本题过程繁多，持续时间也很长，需要同学们求解时保持清醒，尤其是算 FOLLOW 集时更要打起十二分精神。由于给出了详细过程，本文档篇幅较长。

需要指出的是，本文档采用孙景昊式解法：对于每一个非终结符，默认其 FIRST 集不含 ε ，FOLLOW 集不含 $\#$ ，可能与《编译原理》课本的结果有所不同。