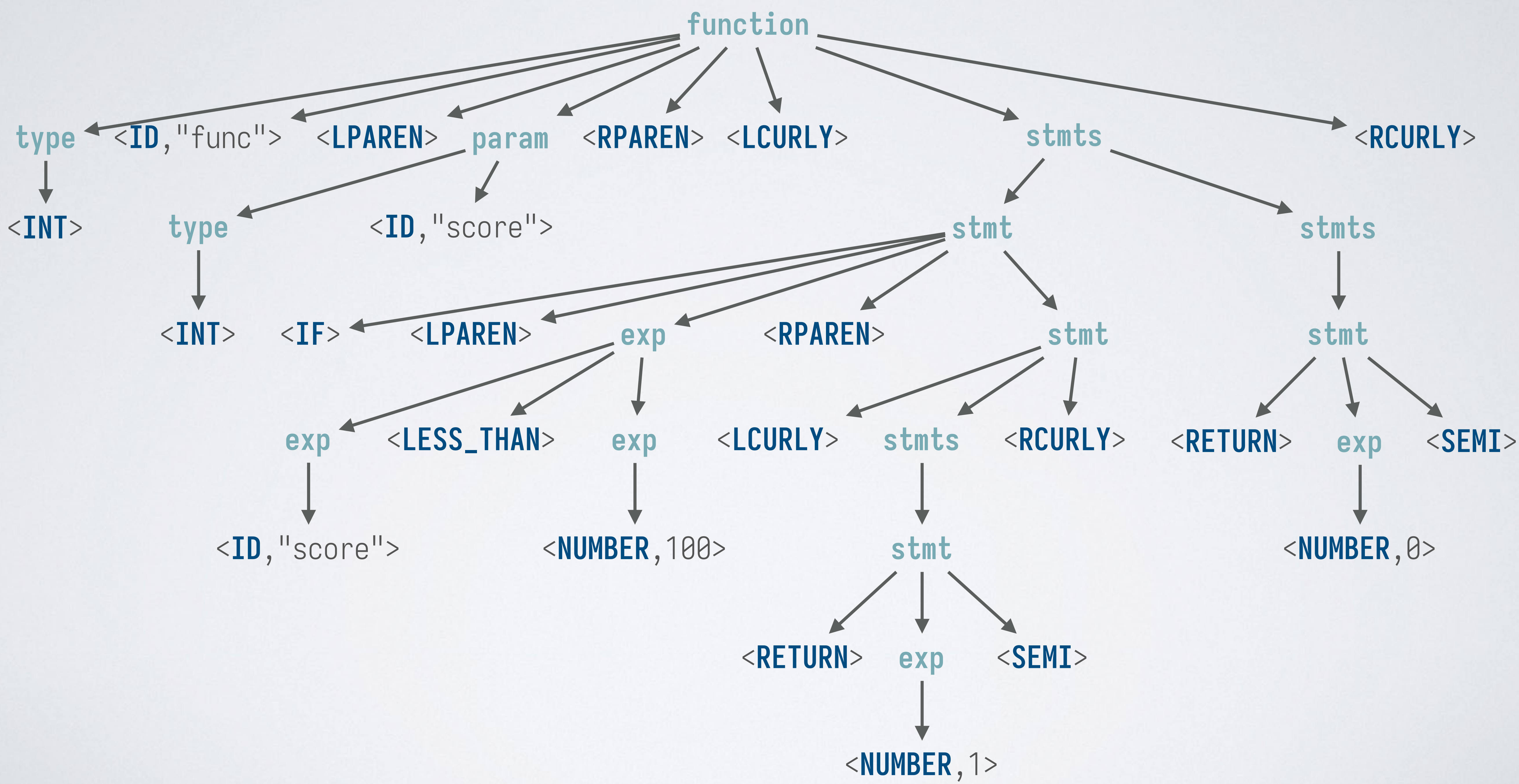
参考答案（0305）

- 1. 语法分析器处理 token 流，输出语法分析树。



参考答案（0305）

- **2. 考虑上下文无关文法 $G = (V_T, V_N, S, P)$：**

  $V_T = \{+, *, \mathrm{a}\}, V_N = \{S\}, P$ 中规则表示为 $S ::= S\,S + \mid S\,S * \mid \mathrm{a}$

  针对符号串 a a a + *, 回答以下问题:

  (1) 给出这个串的一个最左推导

$$
\begin{aligned}
S &\Rightarrow S\,S\,* \\
&\Rightarrow \mathrm{a}\,S\,* \\
&\Rightarrow \mathrm{a}\,S\,S + * \\
&\Rightarrow \mathrm{a\,a}\,S + * \\
&\Rightarrow \mathrm{a\,a\,a} + *
\end{aligned}
$$

参考答案（0305）

- **2. 考虑上下文无关文法 $G = (V_T, V_N, S, P)$：**

$V_T = \{+, *, \mathrm{a}\}, V_N = \{S\}, P$ 中规则表示为 $S ::= S\,S + \mid S\,S * \mid \mathrm{a}$

针对符号串 a a a + *, 回答以下问题:

(2)给出这个串的一个最右推导

$$
\begin{aligned}
S &\Rightarrow S\,S\,* \\
&\Rightarrow S\,S\,S + * \\
&\Rightarrow S\,S\,\mathrm{a} + * \\
&\Rightarrow S\,\mathrm{a}\,\mathrm{a} + * \\
&\Rightarrow \mathrm{a}\,\mathrm{a}\,\mathrm{a} + *
\end{aligned}
$$

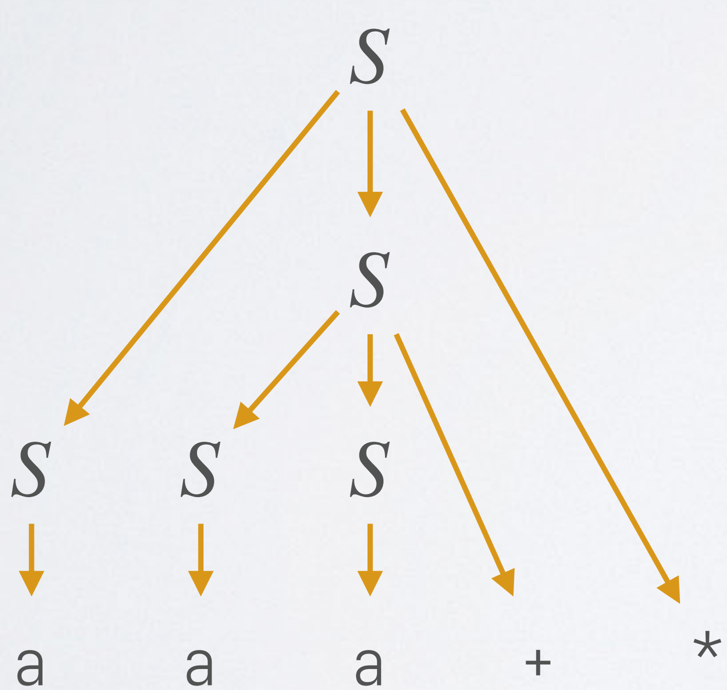
参考答案（0305）

- **2. 考虑上下文无关文法 $G = (V_T, V_N, S, P)$：**

  $V_T = \{+, *, \mathrm{a}\}$, $V_N = \{S\}$, P 中规则表示为 $S ::= S\,S + \mid S\,S * \mid \mathrm{a}$

  针对符号串 a a a + *, 回答以下问题:

  (3) 画出这个串的一棵语法分析树



参考答案（0305）

- **2. 考虑上下文无关文法 $G = (V_T, V_N, S, P)$：**

  $V_T = \{+, *, \text{a}\}, V_N = \{S\}, P$ 中规则表示为 $S ::= S\,S + \mid S\,S * \mid \text{a}$

  针对符号串 a a a + *，回答以下问题：

  (4) 判断这个文法是否是二义性的，并解释你的判断

无二义性。该文法表示后缀表达式，总能够根据串最右侧的符号是 +、* 或 a 来确定唯一的产生规则。

参考答案（0305）

- **3. 为下列的语言设计上下文无关文法：**

  (1) 所有由 a 和 b 组成的形如 $\mathsf{a}^k\mathsf{b}^k$ $(k \geq 0)$ 的符号串

$$\boxed{S ::= \epsilon \mid \mathsf{a}\,S\,\mathsf{b}}$$

证明该文法 G 是正确的，即 $L(G) = L_o$，用 L_o 表示题目要求的语言。

- 证明 $L(G) \subseteq L_o$，即对任意 $S \Rightarrow^* w$ (w 是终结符号串)，有 $w \in L_o$。

  按照推导的步数 k 进行归纳。若 $k = 0$，有 $w = \epsilon \in L_o$。若 $k > 0$，则有 $S \Rightarrow \mathsf{a}\,S\,\mathsf{b} \Rightarrow^* w$，中间的 S 的推导步数为 $k - 1$，根据归纳假设可得 $w = \mathsf{a}\,\underline{\mathsf{a}^n\mathsf{b}^n}\,\mathsf{b}$ (存在自然数 n)，从而 $w \in L_o$。

- 证明 $L_o \subseteq L(G)$，即对任意 $w \in L_o$，存在推导 $S \Rightarrow^* w$。

  按照 w 的长度进行归纳。若 $|w| = 0$，有 $S \Rightarrow \epsilon$。若 $|w| > 0$，那么存在 $k > 0$ 使得 $w = \mathsf{a}^k\mathsf{b}^k = \mathsf{a}\,\underline{\mathsf{a}^{k-1}\mathsf{b}^{k-1}}\,\mathsf{b}$，而根据归纳假设可得 $S \Rightarrow^* \mathsf{a}^{k-1}\mathsf{b}^{k-1}$，从而 $S \Rightarrow \mathsf{a}\,S\,\mathsf{b} \Rightarrow^* w$。

参考答案（0305）

- **3. 为下列的语言设计上下文无关文法：**

  (2) 所有由 a 和 b 组成的回文符号串，即反转后等于自身的符号串

$$\boxed{S ::= \epsilon \mid \mathrm{a} \mid \mathrm{b} \mid \mathrm{a}\,S\,\mathrm{a} \mid \mathrm{b}\,S\,\mathrm{b}}$$

证明该文法 G 是正确的，即 $L(G) = L_o$，用 L_o 表示题目要求的语言。

- 证明 $L(G) \subseteq L_o$，即对任意 $S \Rightarrow^* w$（w 是终结符号串），有 $w \in L_o$。略。
- 证明 $L_o \subseteq L(G)$，即对任意 $w \in L_o$，存在推导 $S \Rightarrow^* w$。

  按照 w 的长度进行归纳。若 $|w| = 0$，有 $S \Rightarrow \epsilon$。若 $|w| = 1$，那么要么 $w = \mathrm{a}$，要么 $w = \mathrm{b}$，两种情况均有 $S \Rightarrow w$。若 $|w| > 1$，由于 w 是回文串，那么要么 $w = \mathrm{a}\,w'\mathrm{a}$，要么 $w = \mathrm{b}\,w'\mathrm{b}$，且 w' 也是回文串，则根据归纳假设，有 $S \Rightarrow^* w'$，从而可得 $S \Rightarrow \mathrm{a}\,S\,\mathrm{a} \Rightarrow^* w$ 或 $S \Rightarrow \mathrm{b}\,S\,\mathrm{b} \Rightarrow^* w$。

参考答案（0305）

- **3. 为下列的语言设计上下文无关文法：**

  (3) 所有由 a 和 b 组成的 a 的数目和 b 的数目相同的符号串

$$\boxed{S ::= \epsilon \mid \mathrm{a}\,S\,\mathrm{b}\,S \mid \mathrm{b}\,S\,\mathrm{a}\,S}$$

证明该文法 G 是正确的，即 $L(G) = L_o$，用 L_o 表示题目要求的语言。

- 证明 $L(G) \subseteq L_o$，即对任意 $S \Rightarrow^* w$（w 是终结符号串），有 $w \in L_o$。略。
- 证明 $L_o \subseteq L(G)$，即对任意 $w \in L_o$，存在推导 $S \Rightarrow^* w$。

  按照 w 的长度进行归纳。若 $|w| = 0$，有 $S \Rightarrow \epsilon$。若 $|w| = k > 0$，令 c_i 表示前缀 $w_1 w_2 \ldots w_i$ 中 a 的数目减去 b 的数目，则 $c_0 = c_k = 0$。

  若 $w_1 = \mathrm{a}$，那么 $c_1 = 1$，且一定存在 i 使得 $c_i = 0$，令 i_o 为其中最小的一个。也就是说，从 c_1 到 c_{i_o-1} 都是正数，那么 w_{i_o} 一定是 b。更进一步，因为 $1 < i_o \leq k$，我们知道子串 $w_2 \ldots w_{i_o-1}$、$w_{i_o+1} \ldots w_k$ 是良定义的（允许为空），且这两个子串中 a 和 b 的数目相同。根据归纳假设，可得 $S \Rightarrow^* w_2 \ldots w_{i_o-1}$ 和 $S \Rightarrow^* w_{i_o+1} \ldots w_k$，故 $S \Rightarrow \mathrm{a}\,S\,\mathrm{b}\,S \Rightarrow^* w$。

  若 $w_1 = \mathrm{b}$，也可用类似的方法证明。

参考答案（0305）

- 3. 为下列的语言设计上下文无关文法：

(4) 所有由 a 和 b 组成的 a 的数目和 b 的数目不相同的符号串

$$
\begin{array}{ll}
S ::= A \mid B & A ::= T\,\mathrm{a}\,T \mid A\,\mathrm{a}\,T \\
B ::= T\,\mathrm{b}\,T \mid B\,\mathrm{b}\,T & T ::= \epsilon \mid \mathrm{a}\,T\,\mathrm{b}\,T \mid \mathrm{b}\,T\,\mathrm{a}\,T
\end{array}
$$

A 表示 a 的数目比 b 多的串，B 表示 b 的数目比 a 多的串，T 表示 a 和 b 的数目相同串。

证明 $L(A) = L_a$，用 $L(A)$ 表示 A 推导的串的集合，用 L_a 表示由 a 的数目比 b 多的串构成的语言。

- 证明 $L(A) \subseteq L_a$，即对任意 $A \Rightarrow^* w$（w 是终结符号串），有 $w \in L_a$。略。
- 证明 $L_a \subseteq L(A)$，即对任意 $w \in L_a$，存在推导 $A \Rightarrow^* w$。

  按照 a 比 b 多的数目 k 进行归纳。令 c_i 表示前缀 $w_1w_2\ldots w_i$ 中 a 的数目减去 b 的数目。则 $c_0 = 0$，$c_n = k$（n 是 w 的长度）。那么一定存在 i 使得 $c_i = k - 1$，令 i_o 为其中最大的一个，则一定有 $w_{i_o+1} = \mathrm{a}$。此时，$c_{i_o+1} = k$，从而子串 $w_{i_o+1}\ldots w_n$ 中 a 和 b 的数目相同，即可以被 T 推导出来。而前缀 $w_1\ldots w_{i_o}$ 中 a 的数目比 b 多 $k-1$，若 $k = 1$，则对应 $A \Rightarrow T\,\mathrm{a}\,T$；若 $k > 1$，则对应 $A \Rightarrow A\,\mathrm{a}\,T$，可以通过归纳完成证明。

参考答案（0305）

- 4. 考虑上下文无关文法，终结符号集合为 {a, b, c, d, EOF}：

$$
\begin{aligned}
S' &::= L\ \mathrm{EOF} \\
L &::= R\ \mathrm{a} \mid Q\ \mathrm{b\ a} \\
R &::= \mathrm{a\ b\ a} \mid \mathrm{c\ a\ b\ a} \mid R\ \mathrm{b\ c} \\
Q &::= \mathrm{b\ b\ c} \mid \mathrm{b\ c}
\end{aligned}
$$

(1) 计算非终结符号 S'、L、R、Q 的 FIRST 和 FOLLOW 集合

$$
\begin{aligned}
\mathrm{FIRST}(S') &= \{\mathrm{a, b, c}\} \\
\mathrm{FIRST}(L) &= \{\mathrm{a, b, c}\} \\
\mathrm{FIRST}(R) &= \{\mathrm{a, c}\} \\
\mathrm{FIRST}(Q) &= \{\mathrm{b}\}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{FOLLOW}(S') &= \{\,\} \\
\mathrm{FOLLOW}(L) &= \{\mathrm{EOF}\} \\
\mathrm{FOLLOW}(R) &= \{\mathrm{a, b}\} \\
\mathrm{FOLLOW}(Q) &= \{\mathrm{b}\}
\end{aligned}
$$

参考答案（0305）

- **4. 考虑上下文无关文法，终结符号集合为 {a, b, c, d, EOF}：**

$$
\begin{aligned}
S' &::= L \ \mathrm{EOF} \\
L &::= R\ \mathrm{a} \mid Q\ \mathrm{b\ a} \\
R &::= \mathrm{a\ b\ a} \mid \mathrm{c\ a\ b\ a} \mid R\ \mathrm{b\ c} \\
Q &::= \mathrm{b\ b\ c} \mid \mathrm{b\ c}
\end{aligned}
$$

(2) 判断该文法是否是 LL(1) 文法，如果不是的话则通过消除左递归、提取左公因子等方法把它转换为一个 LL(1) 文法

上面的文法不是 LL(1) 的，比如 R 有左递归，Q 有左公因子。

$$
\begin{aligned}
S' &::= L \ \mathrm{EOF} \\
L &::= R\ \mathrm{a} \mid Q\ \mathrm{b\ a} \\
R &::= \mathrm{a\ b\ a}\ R' \mid \mathrm{c\ a\ b\ a}\ R' \\
R' &::= \epsilon \mid \mathrm{b\ c}\ R' \\
Q &::= \mathrm{b}\ Q' \\
Q' &::= \mathrm{b\ c} \mid \mathrm{c}
\end{aligned}
$$