

编译原理与技术第一次作业

2026 年 9 月 21 日

作业说明

- 作业内容部分摘自编译原理 (第三版). 陈意云, 张昱. 高等教育出版社. 2014.
- 没有序号的题目为原创题。
- 本次作业只接受电子版提交。完成所有内容后，请将你的作业命名为“学号 _ 姓名 _HW1.pdf”并上传至希冀平台。
- 本次作业 DDL 为 9 月 28 日 23:59，迟交将被扣除一定比例的分数。

作业正文

2.3 叙述由下列正规式描述的语言。

(d) $0^*10^*10^*10^*$

2.4 为下列语言写出正规定义。

(e) 最多只有一处相邻数字相同的所有数字串

(g) 由偶数个 0 和奇数个 1 构成的所有 0 和 1 的串

(h) 所有不含子串 011 的 0 和 1 的串

2.7 用算法 2.4 (原书 P32) 为下列正规式构造不确定有限自动机, 给出处理输入串 *ababbab* 的状态转换序列。

(b) $(a^* | b^*)^*$

2.12 为下列正规式构造最简 DFA。

(a) $(a | b)^* a (a | b)$

(b) $(a | b)^* a (a | b) (a | b)$

正规式与最简 DFA

设字母表 $\Sigma = \{a, b\}$ 。子串的一次出现由其起始位置确定。

构造一个正规式, 描述所有恰好包含一次子串 *ab* 的字符串; 并为该语言构造最简 DFA, 给出完整的状态转换表。最后, 分别给出该 DFA 处理输入串 *bbaabbaa* 和 *abab* 时的状态转换序列, 并判断它们是否被接受。

从正规式到词法分析器

Cminusf 的词法分析器需要识别标识符和数值。设

$$\begin{aligned} \text{letter} &= a | \cdots | z | A | \cdots | Z, \\ \text{digit} &= 0 | \cdots | 9. \end{aligned}$$

其中, 标识符 ID 是由一个或多个字母组成的字符串; 整数 INTEGER 是由一个或多个数字组成的字符串; 浮点数 FLOATPOINT 由数字和恰好一个小数点组成, 小数点两侧至少一侧非空。例如 12.、.5 和 12.5 均为合法浮点数, 而单独的 . 不是合法浮点数。

- (a) 为 FLOATPOINT 写出正规定义。
- (b) 构造识别 FLOATPOINT 的最简 DFA。可以将任意数字统一记作输入类别 d 。请标明初态、接受状态和陷阱状态，并给出完整的状态转换表。最后，分别给出该 DFA 处理输入串 12.5 和 . 时的状态转换序列，并判断它们能否作为一个完整的 FLOATPOINT 被接受。