

Exam-M0406 字符串

一、实现一个简易的样式匹配和替换工具 my_sed

类似于 UNIX 下的 sed，具体要求如下：

1. 搜索样式中支持如下正则表达式的元字符：

1. `.` 匹配除新行符 (`'\n'`) 之外的任意单个字符
2. `[]` 匹配或不匹配 (`[^]`) 出现在 `[]` 中的任意单个字符
3. `?` 匹配?之前的子表达式 (或字符) 0~1 次
4. `*` 匹配*之前的子表达式 (或字符) 0~多次
5. `+` 匹配+之前的子表达式 (或字符) 1~多次
6. `{n}` 匹配其前面的子表达式 (或字符) n 次 (n 为正整数)
7. `{n,}` 匹配其前面的子表达式 (或字符) 最少 n 次 (n 为正整数)
8. `{n,m}` 匹配其前面的子表达式 (或字符) 最少 n 次, 最多 m 次 (n 、 m 为正整数)
9. `^` 锚定行首
10. `$` 锚定行尾
11. `\<` 和 `\>` 锚定单词前后边界
12. `\b` 锚定单词前后边界
13. `\B` 锚定非单词边界
14. `()` 子表达式
15. `\n` 反向引用 (back reference) 第 n 个子表达式匹配的结果 (n 为正整数)

2. 替换串样式支持如下元字符：

1. `&` 将搜索串匹配的内容作为替换串的一部分
2. `\n` 将搜索串中第 n 个子表达式匹配的内容作为替换串的一部分
3. `\U` 将随后的字母全部换成大写
4. `\u` 将随后的第一个字母换成大写
5. `\L` 将随后的字母全部换成小写
6. `\l` 将随后的第一个字母换成小写
7. `\E` 停止 `\U`、`\L` 的影响

3. 命令执行：

1. 命令格式：

1. 搜索且替换: `my_sed 's/pattern/replaced/'`
2. 仅搜索: `my_sed '/pattern/'`
2. 支持命令选项(仅要求支持一个选项)
 1. `-n` 不打印
3. 指令
 1. `p` print 结果, 如: `my_sed -n '/pattern/p'` #仅打印匹配的行
 2. `d` 删除匹配的行, 如: `my_sed '/pattern/d'`
4. 匹配行为控制
 1. `g` 处理一行中所有匹配的内容, 如: `my_sed 's/pattern/replaced/g'`
 2. `i` 匹配过程中不区分大小写, 如: `my_sed 's/pattern/replaced/gi'`
5. 支持从标准输入和文件读取内容进行匹配
6. 如果输入内容来自文件, 则支持在指定范围进行匹配, 如:
`my_sed '2,12 s/pattern/replaced/'`