

# 实 验 报 告

评分：

计算机科学与技术 系 2021 级      学号 PB21111728      姓名 刘芷辰      日期 2022 年 12 月 7 日

---

实验题目：最长重复子字符串

实验目的：本次实验需要用汇编语言写一个程序，程序的功能是得到一个字符串中重复子字符串的最大长度，并将结果存储在内存中。具体要求如下：

- 程序应当从 x3000 开始运行，程序的功能是读取预先存放在 x3100 的 N，x3101 开始的连续空间的字符串 S，得到最长重复子字符串的长度，并在程序运行结束后，将所要求的结果保存在 x3050
- 约定：程序运行到空指令（对于未初始化，也没有在程序运行过程中修改的内存，认为它存放一条空指令）或 TRAP 指令时，运行结束
- 实验每次需要提交相应的机器码或汇编代码
  - 对于本次实验，提交汇编代码
- 汇编代码注意格式：
  - 需要以 .ORIG x3000 开始
  - 需要以 .END 结束
  - 最后一条指令为 HALT
  - 关键词大写
  - 逗号后需要空格

实验的部分测试数据如下：

# 实 验 报 告

评分:

计算机科学与技术 系 2021 级 学号 PB21111728 姓名 刘芷辰 日期 2022 年 12 月 7 日

Memory address	x3050	...	x3100	x3101	x3102	x3103	x3104	x3105	x3106
example 1	RESULT=3		NUM=6	a	a	b	b	b	c
example 2	RESULT=4		NUM=5	Z	Z	Z	Z	z	
example 3	RESULT=3		NUM=6	a	a	b	a	a	a

实验原理:

代码如下图:

```
;R0读入N, R1指针, R2最大长度, R3当前长度
;R4当前字符, R5上一位字符, R6和R7作tmp
.ORIG    x3000
AND      R0, R0, #0;
AND      R1, R1, #0;
AND      R2, R2, #0;
AND      R3, R3, #0;
AND      R4, R4, #0;
AND      R5, R5, #0;
AND      R6, R6, #0;
AND      R7, R7, #0; 初始化
LDI      R0, NUM;
LD        R1, DATA;
LDR      R4, R1, #0;
ADD      R2, R2, #1;
LABEL2   AND      R3, R3, #0; 当前与上一位不同, 当前长度重新计数
LABEL1   ADD      R3, R3, #1; 当前长度加1
          ADD      R0, R0, #-1;
          BRz      FINISH; 全部读入则跳转

          ADD      R5, R4, #0;
          ADD,     R1, R1, #1;
          LDR      R4, R1, #0; 更新当前位和上一位

;判断前一位和当前位是否相同
          ADD      R6, R4, #0;
          NOT      R6, R6;
          ADD      R6, R6, #1;
          ADD      R7, R5, R6; 得到R7=R5-R4
          BRz      LABEL1; 相同则跳转LABEL1, 不同则比较当前长度和最大长度

;比较当前长度和最大长度
          ADD      R6, R2, #0;
          NOT      R6, R6;
          ADD      R6, R6, #1;
          ADD      R7, R3, R6; 得到R7=R3-R2
          BRn      LABEL2;
          ADD      R2, R3, #0;
```

# 实 验 报 告

评分:

计算机科学与技术 系 2021 级 学号 PB21111728 姓名 刘芷辰 日期 2022 年 12 月 7 日

```
BRnzp LABEL2;当前更长则更新最大长度，两种情况都跳转LABEL2
;最后出现的字符是最长的情况在此作判断
FINISH ADD R6, R2, #0;
      NOT R6, R6;
      ADD R6, R6, #1;
      ADD R7, R3, R6;
      BRn STORE;
      ADD R2, R3, #0;
STORE  STI R2, RESULT;
      HALT

RESULT .FILL x3050
NUM     .FILL x3100
DATA    .FILL x3101

      .END
```

共分为如下三步：

## 1. 数据初始化：

把分别位于内存 x3100 的 N 存入寄存器 R0 中,R1 读入地址 x3101, 用作指针, R2 清零用来储存最大长度, R3 清零用来储存当前重复子字符串的长度, R4 清零用来储存当前字符, R5 清零用来储存上一位字符, R6 和 R7 清零用来作 tmp 便于实现其他操作而不影响其他寄存器储存的值;

## 2. 读取字符作判断

(1) 每次通过指针 R1 加 1 将内存的下一位字符读到 R4 中, 并与 R5 (上一位字符) 中的值作比较, 若二者相同, 则 R3 的值加 1, 若不同, 则表示当前重复子字符串结束, 进行 (2)

# 实 验 报 告

评分:

计算机科学与技术 系 2021 级 学号 PB21111728 姓名 刘芷辰 日期 2022 年 12 月 7 日

(2) 当前重复子字符串结束后, 将 R3 中储存的当前重复子字符串的长度和已经读取过的最大重复子字符串的长度进行比较, 若当前的长度更大, 则更新最大长度

(3) 无论是否更新最大长度, 在经过 (2) 之后, 将 R3 置 0, 再次进行上述操作, 直到所有字符判断完毕

(4) 为解决最长重复子字符串是最后一个重复子字符串的情况 (此时不会在 (1) 中被判断出来), 在所有字符判断完毕之后, 再进行一次

(2) 操作

## 3. 存储结果

把 R3 中的结果通过 ST 指令存在内存 x3102 中

## 实验步骤:

1. 读懂实验目的
2. 将实现过程模块化三个部分: 初始化、读取字符作判断、存储结果
3. 将步骤转为汇编码实现
4. 存为 asm 文件并运行为 obj, 查找错误:

在第一次写完代码运行时发现, 只通过了两个样例, 最后一个没通过, 观察数据发现没通过的样例数据最长重复子字符串出现在末尾, 由此发现对此类情况并不能通过当前位和上一位不同而判别出来, 因此在

# 实 验 报 告

评分:

计算机科学与技术 系 2021 级 学号 PB21111728 姓名 刘芷辰 日期 2022 年 12 月 7 日

结束时再进行一次当前长度和最大长度的比较即可解决此问题

5. 修改完成后在自测网站进行自测

6. 自测通过后书写实验报告

实验结果：(输入数据字符用相应的 `ascii` 码输入)

测试数据 (1):

example 1	RESULT=3	NUM=6	a	a	b	b	b	c
-----------	----------	-------	---	---	---	---	---	---

寄存器:

R0	x0000	0	
R1	x7FFF	32767	
R2	x0003	3	
R3	x0001	1	
R4	x0063	99	
R5	x0062	98	
R6	x2FFE	12286	
R7	xFFFFE	65534	
PSR	x0002	2	CC: Z
PC	x036C	876	
MCR	x0000	0	

内存:

! ▶ x3050 x0003 3

# 实 验 报 告

评分:

计算机科学与技术 系 2021 级 学号 PB21111728 姓名 刘芷辰 日期 2022 年 12 月 7 日

R2 存储结果是 3, 内存 x3050 中存储也为 3, 最长重复子字符串是 bbb, 长度为 3, 结果正确

测试数据 (2):

example 2      RESULT=4      NUM=5    Z    Z    Z    Z    z

寄存器:

R0	x0000	0	
R1	x7FFF	32767	
R2	x0004	4	
R3	x0001	1	
R4	x007A	122	
R5	x005A	90	
R6	x2FFE	12286	
R7	xFFFF	65533	
PSR	x0002	2	CC: Z
PC	x036C	876	
MCR	x0000	0	

内存:

! ▶ x3050      x0004      4

R2 存储结果是 4, 内存 x3050 中存储也为 4, 最长重复子字符串是 ZZZZ, 长度为 4, 结果正确

# 实 验 报 告

评分:

计算机科学与技术 系 2021 级 学号 PB21111728 姓名 刘芷辰 日期 2022 年 12 月 7 日

测试数据 (3):

example 3      RESULT=3      NUM=6      a      a      b      a      a      a

寄存器:

R0	x0000	0	
R1	x7FFF	32767	
R2	x0003	3	
R3	x0003	3	
R4	x0061	97	
R5	x0061	97	
R6	x2FFE	12286	
R7	x0001	1	
PSR	x0002	2	CC: Z
PC	x036C	876	
MCR	x0000	0	

内存:

! ▶ x3050      x0003      3

R2 存储结果是 3, 内存 x3050 中存储也为 3, 最长重复子字符串是 aaa, 长度为 3, 结果正确

# 实 验 报 告

评分:

计算机科学与技术 系 2021 级      学号 PB21111728      姓名 刘芷辰      日期 2022 年 12 月 7 日

自测网站测试结果:

汇编评测

3 / 3 个通过测试用例

- 平均指令数: 86.333333333333
- 通过 6:aabbbc:3, 指令数: 93, 输出: 3
- 通过 5:ZZZZz:4, 指令数: 74, 输出: 4
- 通过 6:aabaaa:3, 指令数: 92, 输出: 3

三个测试样例均正确