

# 评分细则

本课程实验部分分数由最终编译器在测试点上的运行正确性和性能共同决定。

## 测试点

测试点共100个，其中80个为功能测试点，20个为性能测试点，每个测试点1分。

功能和性能测试样例中，均有10个功能测试点，5个功能测试点不公开，作为隐藏点。（隐藏点和非隐藏点分数占比相同）

## 功能测试

使用spike模拟器，运行编译得到的汇编，和 `gcc 9.5.0` 进行对比。

对于功能测试点，只比对输出的字符串，即 `strcmp(src, dst) == 0` 则判定为得分否则不得分

## 性能测试

使用spike模拟器，对汇编实际运行的指令条数进行统计，并求出周期数

指令类型	周期数
转跳	1.2
load/store	4
mul	10
div	100
普通指令(ALU等)	1

例如如果一个汇编代码模拟执行，包含100条add，20条转跳，5条乘法，1条除法，则

$$100 * 1 + 20 * 1.2 + 5 * 10 + 100 = 274$$

每个性能测试点，设置三档得分0.5 1.0

具体评分方式如下，还以上述测试为例：

课程组会综合考虑GCC, Clang等领域通用编译器，为每一个测试点设置三档，例如对于上述测试用例为

指令周期数	得分
500	0.5
200	1.0

设置评分点的周期数为 $pos_{1.0}$

$x$ 为同学运行实际的周期数

$$score = \begin{cases} 0.5 & x \in [pos_{0.5}, INF] \\ \sqrt{\frac{pos_{0.5}-x}{pos_{0.5}-pos_{1.0}}} * 0.5 + 0.5 & x \in [pos_{1.0}, pos_{0.5}] \\ 1.0 & x \in (0, pos_{1.0}] \end{cases}$$

则上述同学得分为:

$$\sqrt{\frac{500 - 274}{500 - 200}} * 0.5 + 0.5 = 0.934$$

由于选择凸函数作为评分标准, 即使周期数为400, 也可以得到

$$\sqrt{\frac{500 - 400}{500 - 200}} * 0.5 + 0.5 = 0.789$$