# 编译原理课程实验（2014年）

在编译原理与技术课程的大作业中，每个小组将独立完成一个语言的相关编译器的开发。你们将建立一个1~3人的项目小组，合作完成大作业。

## 第一阶段

### 设计语言

参照我们提供的示例代码，设计语言MyLang。要求设计的语言可以描述示例代码中的功能。

设计的语言要求支持常见计算表达式（例如+，-，\*，/，%，括号，优先级处理等）。

MyLang是强类型语言，关于程序语言的类型，请参考下面一些概念：

* 静态类型定义语言：一种在编译时，数据类型是固定的语言。大多数静态类型定义语言强制这一点，它要求你在使用所有变量之前要声明它们的数据类型。Java和C是静态类型定义语言。
* 动态类型定义语言：一种在执行期间才去发现数据类型的语言，与静态类型定义相反。VBScript和Python是动态类型定义的，因为它们是在第一次给一个变量赋值的时候找出它的类型的。
* 强类型定义语言：一种总是强制类型定义的语言。Java和Python是强制类型定义的。如果你有一个整数，如果不显式地进行转换，你不能将其视为一个字符串。
* 弱类型定义语言：一种类型可以被忽略的语言，与强类型定义相反。VBScript是弱类型定义的。在VBScript中，可以将字符串'12'和整数3进行连接得到字符串'123'，然后可以把它看成整数123，而不需要显示转换。

除我们的示例程序中用到的功能，你还可以在我们提供的程序的基础上扩充你的语言的特性。

### 熟悉语言

使用MyLang编写快速排序程序，以熟悉MyLang语言。

### 书写BNF

使用BNF定义MyLang语言，以PDF文件形式提交。

你可以参考一些网络资料，如：<http://www.daimi.au.dk/dRegAut/JavaBNF.html>

## 第二阶段

### 熟悉汇编

使用C++书写快速排序程序，并将其编译成为汇编代码。

使用汇编程序书写快速排序，并使用汇编器执行。

### 前端实现

为你的编译器提供一个图形界面，以很清楚地显示四个主要阶段的输入和输出（例如，提供四个菜单按钮，分别对应词法分析、语法分析、语义检查、代码生成）。四个阶段的输入输出均需要能够独立显示。

你可以基于JavaCC或者其他类似的编译器生成器开发，也可以自己手工开发。

书写词法分析，生成Token Stream。如果程序违反词法规则，则报出合适的错误信息；

书写语法分析，生成一个可展示的抽象语法树结构。如果程序违反语法规则，则报出合适的错误信息；

## 第三阶段

### 编译器后端

书写语义检查，显示所有变量的类型，及检查其正确性。如果类型不正确，则予以报错，注意类型的兼容性，及函数的形参与实参兼容；

书写代码生成，生成汇编代码。试运行你书写的快速排序程序的汇编代码，如果自动生成后需要手动修改汇编代码才能使其运行的话，修改量不能超过生成代码量的10%。

### 进阶功能

实现其他更有特色，更复杂的功能。

使用MyLang书写一个稍复杂的程序，如八皇后问题，并编译。

### 文档

项目最终提交的时候需要提供附录文件，包含相应编译器说明、语法规则、测试用例等。