# 编译原理实践第 2课 --- 正则表达式与NFA

说明:

(1) 本次作业提交最晚时间为10月10日晚23:59。

(2) 将您的解答填写至每道题后. 如果是程序题, 代码量不大的话，请将代码放至文本框内，如下所示:

#/usr/bin/perl

print "HELLO WORLD!\n”

代码量大的话, 请提交代码文件, 并在相应地方填写，如下所示:

见\*\*\*.pl文件

(3) 如果有多个提交文件, 请打包压缩后提交. 请使用zip、tar.gz等压缩格式，请匆使用rar格式。

一、正则表达式与Latex

以下1. 和3.不需要作答, 只需要回答2.

1. **什么是LaTex？**

LaTeX是一种基于ΤΕΧ的排版系统，由美国计算机学家莱斯利·兰伯特（Leslie Lamport）在20世纪80年代初期开发，利用这种格式，即使使用者没有排版和程序设计的知识也可以充分发挥由TeX所提供的强大功能，能在几天，甚至几小时内生成很多具有书籍质量的印刷品。对于生成复杂表格和数学公式，这一点表现得尤为突出。因此它非常适用于生成高印刷质量的科技和数学类文档。这个系统同样适用于生成从简单的信件到完整书籍的所有其他种类的文档。

1. **从Latex到PDF**

将LaTex源文件解析的结果构建为pyfPDF能够识别格式比如HTML，并通过pyfPDF生成pdf（example.pdf）

正则表达式

Latex源文件

HTML文件

PDF

FPDF

example.tex的标记很多，本次实验不要求完全实现，尽量能解析多的标记即可，但至少要完结解析如下标记：

* document
* title
* abstract
* section
* subsection
* itemize/item
* tabular
* emph
* textbf

**请给出你的程序，该程序读取输入文件example.tex, 输出以上各字段的值（如果有值的话），每个字段可能出现多次，将有多个值。**

**答:**

1. **pyFPDF**

一种将网页转换为pdf的工具

项目主页：https://github.com/reingart/pyfpdf

安装方法：

1. 在线安装：pip install fpdf
2. 本地安装：python setup.py install

教程：https://github.com/reingart/pyfpdf

二、编程题

(1) 正规表达式转NFA

读取input.txt文件, 该文件中给出待转换的正则表达式, 将转换的NFA输出至文件output.txt.

例如, input.txt的内容为如下:

ab\*(a\*|(ab)\*)

输出的output.txt为:

2--[-]-->3

3--[-]-->2

3--[-]-->5

5--[-]-->3

5--[-]-->4

3--[a]-->5

8--[-]-->6

7--[b]-->8

3--[a]-->7

3--[-]-->6

6--[-]-->3

6--[-]-->4

9--[-]-->4

3--[b]-->9

3--[-]-->4

4--[-]-->3

4--[b]-->10

2--[b]-->3

0--[a]-->2

注: 1. 打印的是NFA图中各条边的信息. 如第一条边, 表示节点2经过符号-到达节点3. 为了方便起见, 用符号-表示空串ε。

2. 不同的实现方法，得到的NFA不一致, 所以你得到的可能是另外一个结果。

答: