上课讲义

# 课前的话

刚刚讲完神学—设计模式。接下来由我给大家带来curl 和 正则表达式的讲解。

## 在讲解今天的知识前，我们要先提一个东西

## 爬虫，

那爬虫是什么呢? **网络爬虫是是一种按照一定的规则,自动的抓取万维网信息的程序或者脚本**

**。**

**从中我们可以发现，爬虫需要对信息进行抓取并使用一定的规则进行匹配。**

**刚好我们今天学习的东西就可以实现这个**

**首先**

# curl

**首先,肯定有人会问curl是什么东西呢。cUrl是一个库，他集成了我们模拟一次请求所需要的一系列功能。**

**具体参数可以在文档中查阅～～这里不细讲了**

**3正则表达式**

**说完curl，接下来我们讲讲正则表达式。跟上面一样，肯定有人会问，正则表达式是什么鬼，我们为啥要学他？**

**首先回答第一个问题**

**正则表达式是一种表示方式，让你可以查找匹配特定准则的文本，**换句话说，正则表达式就是记录文本规则的代码**例如，“以字母a开头的”。此表示方法让你可以写一个表达式，选定或匹配多个数据字符串。**

**如果我们掌握好他，我们就可以优雅的处理各种各样的字符串，并从中匹配到我们需要的信息。**

**首先我们从元字符开始学起～～**

**就如上面写的正则表达式代表一类的字符串。所以肯定需要可以代替一类字符的字符。而这种字符我们称作元字符。那我们看看元字符有哪些～**

**有时候我们会要求匹配一个定长的字符串。例如学号～，我们一般是一个10为长的数字，这样我们怎么匹配呢？**

**这时候我们就会用到正则表达式的限定符了。**

**限定符可以表示某个字符出现的次数。**

**我们注意\*他表示匹配任意多个字符。那这任意多个是匹配到哪了？**

**、这是我们就要说关于贪婪和懒惰了～**

**例如 aaaabaaaab**

**我们使用 .\*b 他会匹配到哪里呢？**

**这时候我们想象如果我们要匹配一串相同字符的**

**怎么办呢？**