正则表达式（regur expression）

* 一个正则表达式就是由普通字符以及特殊字符（元字符）组成的文字模式，用于描述字符串的特征。
* 不要希望一下全掌握，只需理解正则表达式能做什么与用法，以后再查便可
* 三个重要的正则表达式命令：

1正则表达式中三种类型的括号：

1）方括号“[ ]”用于匹配字符

2）花括号“{ }”用于匹配字符数量

3）圆括号“（）”用于分组

2插入符号“^”表示正则式的开始

3美元符号“$”表示正则式的结束

* Regex obj = new Regex(“[a-z]{10}”)//表示搜索长度为10的a-z的英文字母
* 元字符：
* . 表示除\n（换行符）以外的任意的单个字符，如：

表达式为：a.b；则aab，axb，a3b，a b，a.b，a b（中间为tab）都等可以与之匹配，但是a（换行了）

b无法表示。

表达式为：a\.b；这时.点被转义为普通字符，只有a.b能匹配

* [ ]表示在字符组中罗列出来的字符，任意取单个，如：

表达式为：a[xyz]b；则只有axb，ayb，azb可以与之匹配

表达式为：a[a-z]b；则中间字符为a-z中任意一个字符即可，区分大小写

表达式可以为：a[.a-zA-Z 王0-9]b；.点出现在[]里面，只表示字符,在[]中\.和.表示一个意思

表达式为：a[a-zA-Z]b；要想实现大小写共用，只能这种写法或者

[A-z]（按照ASCII码顺序），不能写[a-Z]。

表达式为[-a-z]时，开头或结尾的-单纯表达一个字符，不表示范围

* | 表示“或“的意思，优先级最低，一个表达式中可以有多个|

式1：a(x|y)b;匹配axb和ayb

式2：z|food；匹配z或者food

式3：（z|f）ood；与式1一样

* （）表示“改变优先级“和表示”提取组“
* \*表示限定前面单个字符出现n次（n可以为0），想要限定多个用（）

式：a.\*b；此处限定了.出现的次数，此处的点也为元字符，

所以匹配：a任意b（当然不包括换行）

* +表示前面单个字符出现n次（n>=1），想要限定多个用（）

式：a.+b；不匹配ab

* ？表示前面单个字符出现0次或1次，？的另一个意思是终止贪婪模式，(?放在限定次数后面可以终止贪婪模式，一般默认是开启贪婪模式)

想要限定多个用（）

式a.?b；

* { }表示限定前面单个字符出现的次数，想要限定多个用（）

式：[0-9]{8}；可匹配00000000，12345678等

式：[0-9]{4, }；至少4次

式：[0-9]{4,9}；n次（9 >= n >= 4）

* ^表示匹配的字符串必须以^后面的字符或字符串开头，^出现在[ ]

里面时，表示“非“的意思

式：^ab王；

式：a[^abc]b；中括号位置只要不是abc三个字符均可

* $表示匹配的字符必须以$以前的字符结尾

注意：如果一个表达式内没有开始和结尾符号，表示一个字符串中只要有一部分满足正则表达式即可，只有加上开始与结尾符才表示完全限定

* 简写表达式（代替某些元符号的功能）

\d可以表示[0-9],0-9中任意一个字符，区别是\d包含Unicode编译的123，而[0-9]只包含ASCII码的123，而.net默认使用的是Unicode编码，Unicode的123其实就是全角模式下的１２３

\D表示[^0-9]

\w表示用ASCII编码的字符[a-zA-Z0-9\_]，也可以表示中文（特殊），当限定为ASCII匹配时，将不包含中文

\W表示非\w

\s表示所有不可见字符：[space tab \r\n]

\S表示非\s

也可以将他们组合起来就可以表示任意单个字符，如：

[\d\D] [\w\W] [\s\S] 注意.不可表示\n

\b表示单次的边界

* 转换符\

对于元字符，限定符等，如果要匹配字符本身都需要进行转换：\\* \+ \. 等，转换符与字母结合成为简写表达式：\d \w等

正则表达式中将字符串中的\从转义符变为转换符，C#中为@符号。

字符串中@为取消转义

总结：若想在正则表达式中使用转换符号\或者转换符号的功能，前面必须加上@。（不加则为转义功能，使用\时需要多加一个\）

* 贪婪模式（正则表达式默认是开启贪婪模式的）

匹配时，只要还能满足表达式，就一直往后面匹配，直到不能再匹配

如有一个字符串：1111！11111！1111111！111111111！

用正则表达式：@“.+!”； （贪婪模式）匹配时，结果为整个字符串，因为匹配到第一个！后，不满足继续往后面匹配，直到后面没有！为止。

如果用？取消时的表达式：@“.+？！”; (非贪婪模式)，此时匹配到的字符串为1111！（不贪，很稳健！！）