

1.任务描述：合法的IPv4地址由4组数字组成，每组数字之间以.分隔，且取值范围是0-255。请在下列数字中判断出合法的地址。

输入：666666666666

23333333

500.1.2.5

0.0.0.0

192.0.0.1

233.233.233.233

102.52.96.3

In [39]:

```
import re
IPv4_Pattern = re.compile('[0-9]$|[1-9][0-9]$|1[0-9][0-9]$|2[0-4][0-9]$|25[0-5]$') #0-255的字符串正
#枚举了一位数，两位数，1开头的三位数，和2开头的三位数，以及25开头的三位数

IPv4_Address = ['666666666666',
'23333333',
'500.1.2.5',
'0.0.0.0',
'192.0.0.1',
'233.233.233.233',
'102.52.96.3']

for address in IPv4_Address:
    Is_Legal = False #最终判定结果
    result = re.split('[\.]', address) #以.来分隔字符串
    if (len(result) == 4): #首先要有4组数据
        Is_Legal = True #设置结果为True
        for e in result: #对于每组数
            if not (IPv4_Pattern.match(e)): #出现失败匹配，则修改结果为False
                Is_Legal = False
        print(address+" : "+("legal" if Is_Legal else "illgeal"))
```

666666666666 : illgeal

23333333 : illgeal

500.1.2.5 : illgeal

0.0.0.0 : legal

192.0.0.1 : legal

233.233.233.233 : legal

102.52.96.3 : legal

2.任务描述：将句子中的中文日期替换为阿拉伯数字，并改为YYYY-MM-DD格式输出。

输入：新中国于一九四九年十月一日成立。

我的出生日期是二零零零年九月二十三日。

In [1]:

```

import re
number = ["零", "一", "二", "三", "四", "五", "六", "七", "八", "九", "十"] #用下标来映射中文汉字
string = "新中国于一九四九年十月一日成立。 我的出生日期是二零零零年九月二十三日。" #处理文本

#1-4位的年份关键词，年，1-2位的月份关键词，月，一到三位的日关键词，日
pattern = re.compile('[一二三四五六七八九十零]{1,4}年[一二三四五六七八九十]{1,2}月[一二三四五六七八九十]{1,3}日')

result = re.findall(pattern, string) #搜索全部匹配结果

for date in result:
    replacement = ""
    for i in range(len(date)):
        if date[i] in number:
            replacement += str(number.index(date[i])) #通过映射变换
        elif i != len(date)-1:
            replacement += "-"
    num = replacement.split("-")
    replacement = num[0] + "-"
    if len(num[1]) == 3: #对于十一，会映射成101，十二映射成102，因此做特判处理
        replacement += num[1][0] + num[1][2]
    else:
        replacement += num[1]
    replacement += "-"
    if len(num[2]) > 2: #同上，对11-31做特判处理
        replacement += num[2][0] + num[2][-1]
    else:
        replacement += num[2]

    string = re.sub(date, replacement, string)
print(string)

```

新中国于1949-10-1成立。 我的出生日期是2000-9-23。

3.任务描述：将句子中的电话号码提取出来，电话号码格式为区号（可选）后接“-”或空格，再加十位数字。

输入：我的电话号码是1234567890，另一个电话号码是+86-4321567890。如果以上都不可用，您可以在16点前致电+91 5645456545与我联系。

In [3]:

```

import re
#处理文本
text = "我的电话号码是1234567890，另一个电话号码是+86-4321567890。如果以上都不可用，您可以在16点前致电+91 5645456545与我联系。"
pattern = re.compile(r'[+]\d*[-\s]\d{10}|\d{10}') #正则表达式，表示带区号的号码，和不带区号的号码
result = re.findall(pattern, text) #全部匹配结果
print(result)

```

```
['1234567890', '+86-4321567890', '+91 5645456545']
```