

- 1、正则匹配中的贪婪匹配是什么意思？ (5 分)
- 2、html 中 td,li 都是什么元素的子节点？<h2>是什么标签？ (5 分)
- 3、请说明 HTTP 协议中，GET 请求和 POST 请求的区别。 (5 分)
- 4、网络爬虫在抓取页面时可能会碰到什么问题导致无法抓取页面，如何解决。 (5 分)
- 5、进程与线程的区别与关系。 (5 分)
- 6、列举至少 3 种以上你知道的数据库名称。它们分别属于关系型还是非关系型数据库？ (15 分)
- 7、使用 shell、Python、Perl、Ruby 等其中的一种语言编写脚本计算 1 叠加到 100 的结果。 (15 分)
- 8、使用任意语言对以下数组进行增量排序。说明算法的时间复杂度。 (15 分)

```
a=[52,11,10,6,-1,2,2,8,8,15]
```

- 9、使用 xpath 或正则表达式或 javascript 截取下面内容 h3 中的链接地址。 (15 分)

```
<html><body>
<div id="2" class="result c-container " data-click="{ 'rsv_bdr':"0","p5":2}"
tpl="se_com_default" srcid="1599">
<h3 class="t">
<a target="_blank"
href="http://www.baidu.com/link?url=uS0o4bUeQXJr4n3vbMyzO6PxdVuF99KFx4l
BnRPL0L1aPHNCZ7UaAoK4eieWl0seiOLL74wwJp7wdWyTojRglK&wd=&eqid=
ada858db0000f79d00000000359c8b941" data-click="{ 'F':778317EA',
'F1':9D73F1E4', 'F2':4CA6DE6B', 'F3':54E5343F, 'T':1506326850',
'y':EFEC3DDE' }">xxx</a>
</h3>
</body>
</html>
```

- 10、使用 python, java, javascript 或其他语言完成 json 字符串格式化。 (15 分)

```
输入
{"name":"Google","location":{"street":"1600 Amphitheatre
Parkway","city":"Mountain
View","state":"California","country":"US"},"employees":[{"name":"Michael","divisi
on":"Engineering"}, {"name":"Laura","division":"HR"}, {"name":"Elise","division":"
```

Marketing"]}]}

输出

```
{
  "employees": [
    {
      "division": "Engineering",
      "name": "Michael"
    },
    {
      "division": "HR",
      "name": "Laura"
    },
    {
      "division": "Marketing",
      "name": "Elise"
    }
  ],
  "location": {
    "country": "US",
    "state": "California",
    "city": "Mountain View",
    "street": "1600 Amphitheatre Parkway"
  },
  "name": "Google"
}
```