

# 数据结构和算法

作者: 小甲鱼

让编程改变世界

Change the world by program



## 字符串

- 以前的计算机刚被发明的時候，主要作用是做一些科学和工程的计算工作，科学家发明计算机的時候压根儿不可能想到后人还可以用来看毛片。
- 刚开始的计算机都是处理数值工作，后来引入了字符串的概念，计算机开始可以处理非数值的概念了（当然原理还是用数值来模拟非数值，通过ASCII码表）。
- 鉴于大家对字符串已经有足够的了解，我们这几节就需要讲点不一样的东西才能满足大家的口味，我们会谈谈“BoyFriend”算法和“看毛片”算法。



# 字符串

- 我们先来研究下“串”这样的数据结构：
  - 定义：串（String）是由零个或多个字符组成的有限序列，又名叫字符串。
  - 一般记为  $s = a_1a_2a_3\ldots a_n$  ( $n \geq 0$ )
  - 串可以是空串，即没有字符，直接由 "" 表示（注意里边没有空格哦~），或者可以用希腊字母  $\Phi$  来表示（读fai，四声）。
- 子串与主串，例如 'FishC' 是 'FishC.com' 的子串，反之则倒过来。



## 字符串

- 关于子串和主串，小甲鱼在做课件搜集资料的时候还发觉一些有意思的：“over”是“lover”的子串，“end”是“friend”的子串，“lie”是“believe”的子串。可见造字者对这世态是看得清清楚楚明明白白真真切切！
- 相比起来中国的汉字比较有内涵，例如讲男女同睡在一张床上，我们可以仅仅只用“非、羽、白、日”四个字来概括。
- 自己发挥想象力哈~





## 字符串的比较

- 字符串比较大小跟传统的数字比较有点差别，很容易我们可以知道2比1要大，可要是“FishC”和“fishc.com”呢？要怎么比较？比长短？比大小？
- 比大小！没错，比的就是字符串里每个字符的ASCII码大小，因为‘F’ = 70 ‘f’ = 102，‘f’ > ‘F’，所以“fishc.com” > “FishC”
- 其实这样的比较大小没有多大意义，字符串的比较我们更重视是否相等！



## 字符串的存储结构

- 字符串的存储结构与线性表相同，也分顺序存储结构和链式存储结构。
- 字符串的顺序存储结构是用一组地址连续的存储单元来存储串中的字符序列的。
- 按照预定义的大小，为每个定义的字符串变量分配一个固定长度的存储区，一般用定长数组来定义。
- 与线性表相似，既然是固定长度的存储区，就存在一个空间分配不灵活的问题，那么会考虑用链式存储结构。



## 字符串的存储结构

- 不同的是字符串我们一般都是连在一起表述的，“断章取义”的情况并不多，所以习惯上我们通常还是会直接定义一个足够长度的存储区来存储的。



## BF算法

- 事实上，BF算法并不是“BoyFriend”算法的意思，他有一个很黄很暴力的原名：Brute Force
- BF算法属于朴素的模式匹配算法，它的核心思想是：
  - 有两个字符串S和T，长度为N和M。首先S[1]和T[1]比较，若相等，则再比较S[2]和T[2]，一直到T[M]为止；若S[1]和T[1]不等，则T向右移动一个字符的位置，再依次进行比较。
  - 该算法最坏情况下要进行 $M*(N-M+1)$ 次比较，时间复杂度为 $O(M*N)$ 。





## BF算法

- 在这里S是主串，T是子串，这种子串的定位操作通常称作串的模式匹配。
- 假设我们要从主串S=“IloveFishC.com”找到T=“FishC”这个子串的位置，按照BF算法，我们需要进行下边的步骤：



# BF算法

2

I	L	o	v	e	F	i	s	h	C	.	c	o	m
---	---	---	---	---	---	---	---	---	---	---	---	---	---

×

F	i	s	h	C
---	---	---	---	---

3

I	L	o	v	e	F	i	s	h	C	.	c	o	m
---	---	---	---	---	---	---	---	---	---	---	---	---	---

×

F	i	s	h	C
---	---	---	---	---



# BF算法

4

I	L	o	v	e	F	i	s	h	C	.	c	o	m
---	---	---	---	---	---	---	---	---	---	---	---	---	---

×

F	i	s	h	C
---	---	---	---	---

5

I	L	o	v	e	F	i	s	h	C	.	c	o	m
---	---	---	---	---	---	---	---	---	---	---	---	---	---

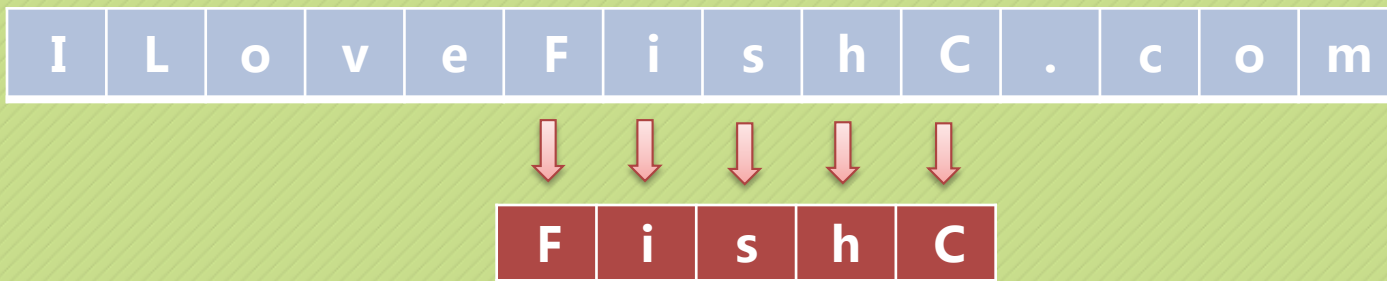
×

F	i	s	h	C
---	---	---	---	---



# BF算法

6



- 课后作业:

- 用你熟悉的一门语言写出BF算法的实现形式, 参考答案详见课件与源代码。
- 考虑为什么BF算法很黄很暴力, 但效率低下?

