

王道考研——数据结构

WWW.CSKAOYAN.COM

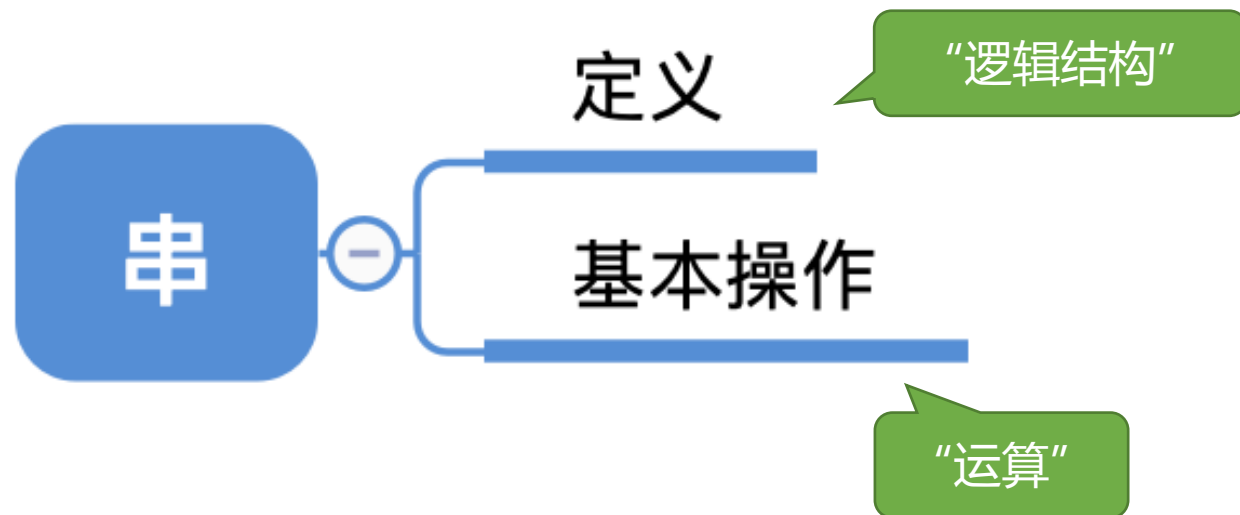
第四章 串

本节内容

串

定义
基本操作

知识总览



注：数据结构三要素——逻辑结构、数据的运算、*存储结构（物理结构）*

存储结构不同，运算的实现方式不同

串的定义

串，即字符串（String）是由零个或多个字符组成的有限序列。一般记为

$S = 'a_1a_2\cdots a_n'$ ($n \geq 0$)

其中， S 是串名，单引号括起来的字符序列是串的值； a_i 可以是字母、数字或其他字符；串中字符的个数 n 称为串的长度。 $n = 0$ 时的串称为空串（用 \emptyset 表示）。

例：

$S = \text{"HelloWorld!"}$

$T = \text{'iPhone 11 Pro Max?'}$

注：有的地方用双引号（如Java、C）
有的地方用单引号（如Python）

子串：串中任意个连续的字符组成的子序列。

主串：包含子串的串。

字符在主串中的位置：字符在串中的序号。

子串在主串中的位置：子串的第一个字符在主串中的位置。

Eg: 'iPhone', 'Pro M' 是串T的子串

Eg: T 是子串'iPhone'的主串

Eg: '1'在T中的位置是8(第一次出现)

Eg: '11 Pro'在T中的位置为8

空串 V.S 空格串：

$M = ''$

M是空串

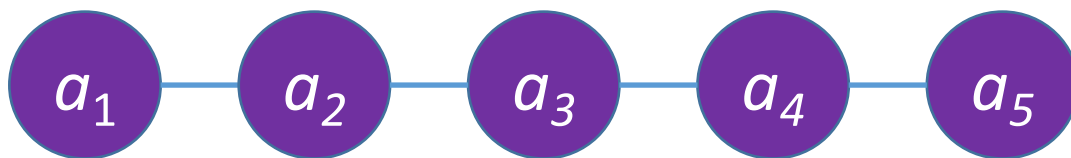
$N = ' \quad '$

N是由三个空格字符组成的空格串，每个空格字符占1B

注意：位序从1开始
而不是从0开始

串 v.s 线性表

串是一种特殊的线性表，数据元素之间呈线性关系



串的数据对象限定为字符集（如中文字符、英文字符、数字字符、标点字符等）

串的基本操作，如增删改查等通常以子串为操作对象



串 V.S 线性表

The image shows two screenshots of Baidu search results. The top screenshot is for the search term 'iphone 11 pro'. It shows the search bar with the text 'iphone 11 pro', tabs for '国内版' (Domestic) and '国际版' (International), and filters for 'ALL', 'IMAGES', and 'VIDEOS'. The results show 51,300,000 results. The first result is 'iPhone 11 Pro - Apple' with a link to 'https://www.apple.com/iphone-11-pro'. The bottom screenshot is for the search term '奥利给'. It shows the search bar with the text '奥利给', tabs for '国内版' and '国际版', and filters for '网页' (Web), '图片' (Image), '视频' (Video), '学术' (Academic), '词典' (Dictionary), and '地图' (Map). The results show 1,530,000 results. The first result is '奥利给_百度百科' with a link to 'https://baike.baidu.com/item/奥利给/23558747'.

国内版 国际版

iphone 11 pro

ALL IMAGES VIDEOS 关闭取词

51,300,000 Results Any time ▾

iPhone 11 Pro - Apple

<https://www.apple.com/iphone-11-pro> ▾

iPhone 11 Pro lets you capture videos that are beautifully true to life, with greater detail and smoother motion. Epic **processing power** means it can shoot 4K video with extended dynamic range and cinematic video stabilization — all at 60 fps. You get more creative control, too, with four times more scene and powerful new editing tools to play with.

国内版 国际版

奥利给

网页 图片 视频 学术 词典 地图

1,530,000 条结果 时间不限 ▾

奥利给_百度百科

奥利给出自快手上的主播在直播或者录视频时的说的话术，该词就是我们常说的“给力”的意思，作为感叹词，可能包含了赞美、加油打气等多种感情色彩。属于网络流行词。

<https://baike.baidu.com/item/奥利给/23558747> ▾

通常以“子串”为增删改查的操作对象

人类的语言通常要多个字符组成的序列才有现实意义

串的基本操作

假设有串 $T=""$ ， $S="iPhone 11 Pro Max?"$ ， $W="Pro"$

StrAssign(&T,chars): 赋值操作。把串 T 赋值为 $chars$ 。

StrCopy(&T,S): 复制操作。由串 S 复制得到串 T 。

StrEmpty(S): 判空操作。若 S 为空串，则返回TRUE，否则返回FALSE。

StrLength(S): 求串长。返回串 S 的元素个数。

ClearString(&S): 清空操作。将 S 清为空串。

DestroyString(&S): 销毁串。将串 S 销毁（回收存储空间）。

Concat(&T,S1,S2): 串联接。用 T 返回由 $S1$ 和 $S2$ 联接而成的新串

SubString(&Sub,S,pos,len): 求子串。用 Sub 返回串 S 的第 pos 个字符起长度为 len 的子串。

Index(S,T): 定位操作。若主串 S 中存在与串 T 值相同的子串，则返回它的主串 S 中第一次出现的位置；否则函数值为0。

StrCompare(S,T): 比较操作。若 $S>T$ ，则返回值 >0 ；若 $S=T$ ，则返回值 $=0$ ；若 $S<T$ ，则返回值 <0 。

Eg: 执行基本操作 **Concat(&T, S, W)** 后， $T="iPhone 11 Pro Max?Pro"$

执行基本操作 **SubString(&T, S, 4, 6)**后， $T="one 11"$

执行基本操作 **Index(S, W)**后，返回值为 11

存储空间扩展？

串的比较操作

StrCompare(S,T): 比较操作。若 $S>T$, 则返回值 >0 ; 若 $S=T$, 则返回值 $=0$; 若 $S<T$, 则返回值 <0 。

A

abandon/ ə'bəndən/ vt. 丢弃; 放弃, 抛弃

aboard/ ə'bo:d/ ad. 在船(车)上; 上船

absolute/ 'æbsəlu:t/ a. 绝对的; 纯粹的

absolutely/ 'æbsəlu:tli/ ad. 完全地; 绝对地

absorb/ əb'sɔ:b/ vt. 吸收; 使专心

abstract/ 'æbstrækt/ n. 摘要

abundant/ ə'bʌndənt/ a. 丰富的; 大量的

abuse/ ə'bjʊ:z, ə'bjʊ:s/ vt. 滥用; 虐待 n. 滥用

academic/ ækə'demik/ a. 学院的; 学术的

accelerate/ æk'seləreit/ vt. (使) 加快; 促进

“abandon” < “aboard”

从第一个字符开始往后依次对比,
先出现更大字符的串就更大

长串的前缀与短串相
同时, 长串更大

“abstract” < “abstraction”

“abstract” < “abstract ”

只有两个串完全相
同时, 才相等

“academic” > “abuse”

“academic” = “academic”

$$y = f(x)$$

字符集：函数定义域

编码：函数映射规则 f

y ：对应的二进制数

任何数据存到计算机中一定是二进制数。

需要确定一个字符和二进制数的对应规则

这就是“编码”

“字符集”：

英文字符——ASCII字符集
中英文——Unicode字符集

基于同一个字符集，可以有多种编码方案，如：UTF-8，UTF-16

注：采用不同的编码方式，每个字符所占空间不同，考研中只需默认每个字符占1B即可

字符集编码

ASCII 字符代码表 一

高四位 低四位		ASCII非打印控制字符										ASCII 打印字符											
		0000					0001					0010	0011		0100	0101		0110		0111			
		0					1					2	3		4	5		6		7			
		+进制	字符	ctrl	代码	字符解释	+进制	字符	ctrl	代码	字符解释	+进制	字符	+进制	字符	+进制	字符	+进制	字符	+进制	字符	ctrl	
0000	0	0	BLANK NULL	^@	NUL	空	16	▶	^P	DLE	数据链路转意	32		48	0	64	@	80	P	96	`	112	p
0001	1	1	☺	^A	SOH	头标开始	17	◀	^Q	DC1	设备控制 1	33	!	49	1	65	A	81	Q	97	a	113	q
0010	2	2	☹	^B	STX	正文开始	18	↕	^R	DC2	设备控制 2	34	"	50	2	66	B	82	R	98	b	114	r
0011	3	3	♥	^C	ETX	正文结束	19	!!	^S	DC3	设备控制 3	35	#	51	3	67	C	83	S	99	c	115	s
0100	4	4	♦	^D	EOT	传输结束	20	¶	^T	DC4	设备控制 4	36	\$	52	4	68	D	84	T	100	d	116	t
0101	5	5	♣	^E	ENQ	查询	21	§	^U	NAK	反确认	37	%	53	5	69	E	85	U	101	e	117	u
0110	6	6	♠	^F	ACK	确认	22	■	^V	SYN	同步空闲	38	&	54	6	70	F	86	V	102	f	118	v
0111	7	7	●	^G	BEL	震铃	23	↑	^W	ETB	传输块结束	39	'	55	7	71	G	87	w	103	g	119	w
1000	8	8	◻	^H	BS	退格	24	↑	^X	CAN	取消	40	(56	8	72	H	88	X	104	h	120	x
1001	9	9	○	^I	TAB	水平制表符	25	↓	^Y	EM	媒体结束	41)	57	9	73	I	89	Y	105	i	121	y
1010	A	10	◻	^J	LF	换行/新行	26	→	^Z	SUB	替换	42	*	58	:	74	J	90	Z	106	j	122	z
1011	B	11	♂	^K	VT	竖直制表符	27	←	^[ESC	转意	43	+	59	;	75	K	91	[107	k	123	{
1100	C	12	♀	^L	FF	换页/新页	28	└	^\ FS	文件分隔符	44	,	60	<	76	L	92	\	108	l	124		
1101	D	13	♪	^M	CR	回车	29	↔	^] GS	组分分隔符	45	-	61	=	77	M	93]	109	m	125	}	
1110	E	14	🎵	^N	SO	移出	30	▲	^6 RS	记录分隔符	46	.	62	>	78	N	94	^	110	n	126	~	
1111	F	15	☼	^O	SI	移入	31	▼	^- US	单元分隔符	47	/	63	?	79	O	95	_	111	o	127	Δ	^Back space

注：表中的ASCII字符可以用：ALT + “小键盘上的数字键”输入

知识回顾与重要考点

串

定义

串，即字符串（String）是由零个或多个字符组成的有限序列

术语：串长、空串、空格串、子串、主串、字符在主串中的位置、子串在主串中的位置

串V.S线性表

串的数据对象限定为字符集

串的基本操作大多以“子串”为操作对象

基本操作

$\text{Index}(S, T)$ ，定位操作，找到串T在主串S中的位置

$\text{StrCompare}(S, T)$ ：比较操作。若 $S > T$ ，则返回值 > 0 ；若 $S = T$ ，则返回值 $= 0$ ；若 $S < T$ ，则返回值 < 0 。

其他...

想想四六级词汇表如何排序

字符集编码

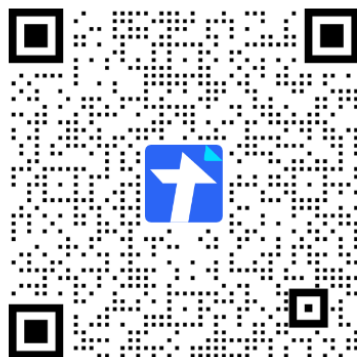
每个字符在计算机中对应一个二进制数，比较字符的大小其实就是比较二进制数的大小

欢迎大家对本节视频进行评价~



学员评分：4.1_1 串的...

扫一扫二维码打开或分享给好友



- 腾讯文档 -

可多人实时在线编辑，权限安全可控



公众号：王道在线



b站：王道计算机教育



抖音：王道计算机考研