HyperionDev

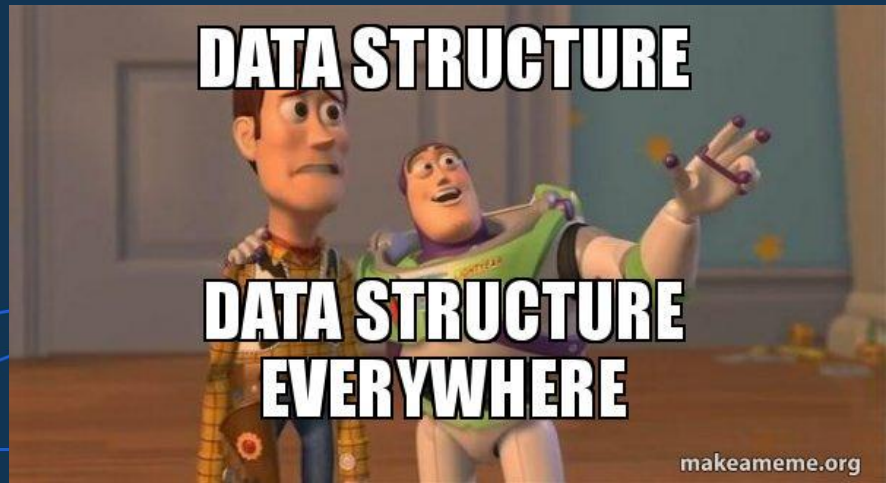
Linear Data Structures II: Strings

Today’s concepts

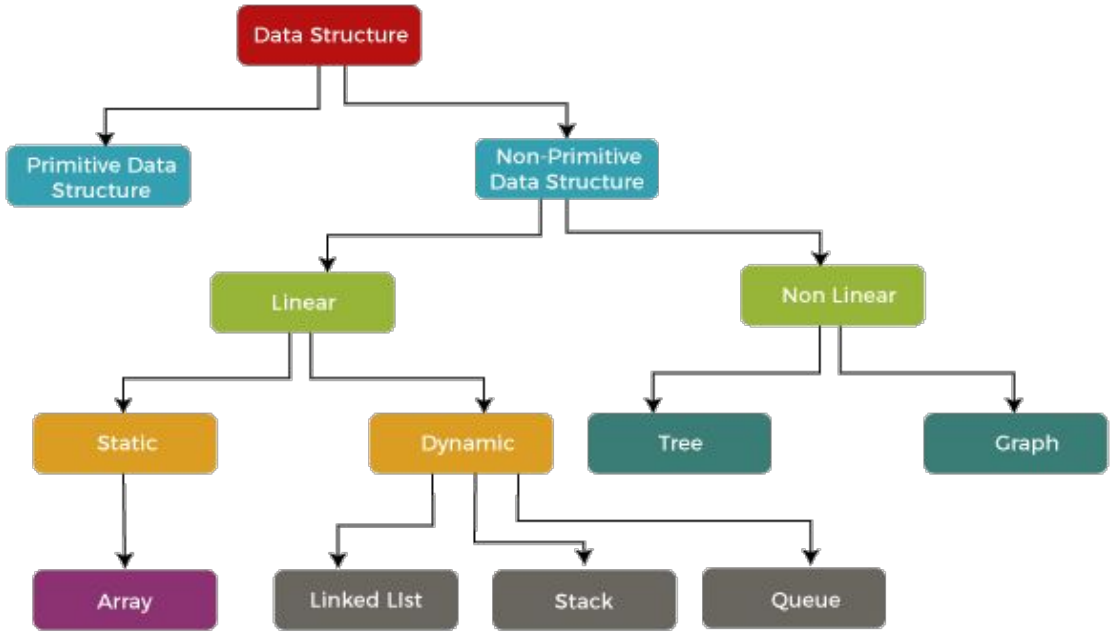
- Recap: Data structure, arrays and linked list
- String data type
- Character encoding: ASCII vs Unicode
- String representations

What is a data structure?

A **data structure** is a particular way of organizing data in a computer so that it can be used effectively.

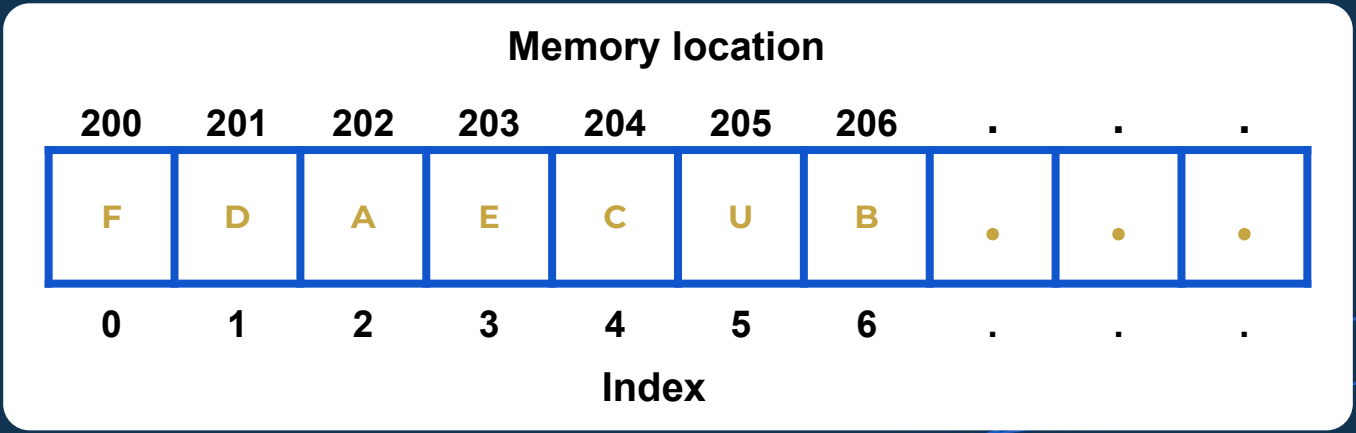


Types of data structures



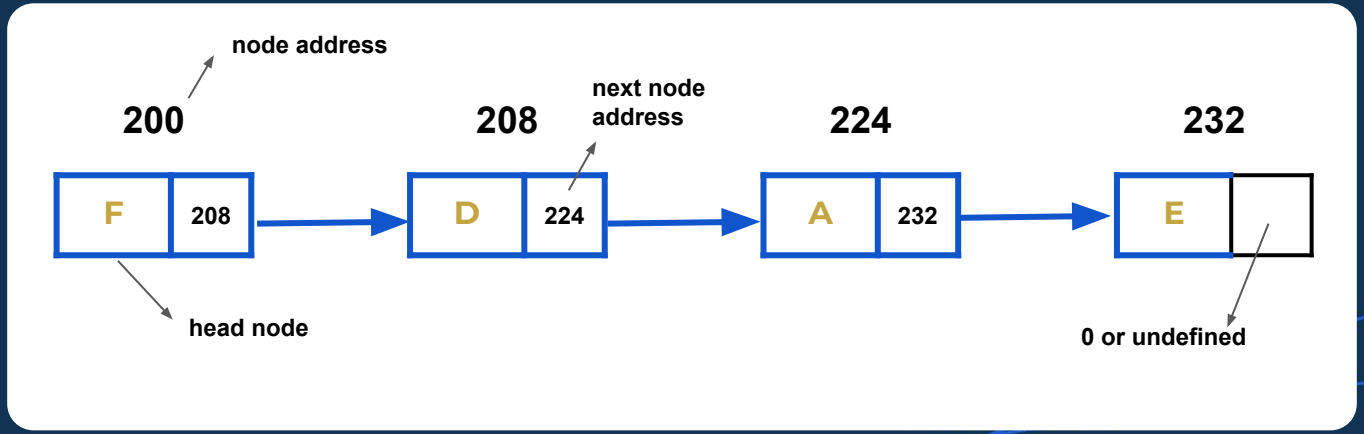
Arrays

An **array** is a collection of items stored at contiguous memory locations.



Linked lists

A **linked list** is a linear data structure, in which the elements are not stored at contiguous memory locations.



String data type

A **string** is a sequence of characters, often used to represent text.

Language	String declaration
Pascal	var myString: string[11] myString = 'Hello World';
C	char myString[] = "Hello World";
C++	string myString = "Hello World";
Java	String myString = "Hello World";
Javascript	let myString = "Hello World";
Python	myString = "Hello World"

Operations performed on Strings

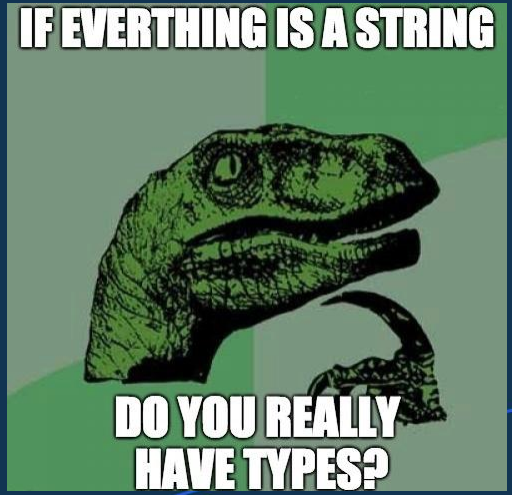
- String concatenation
- Reading and writing strings
- Comparing strings for equality
- Extracting a portion of a string
- Copying one string to another

Advantages of Strings

- Ease of Use
- Compatibility
- Memory Efficiency
- Data Representation: E.g. “YYYY-MM-DD” and “HH:MM:SS”.
- Text Processing

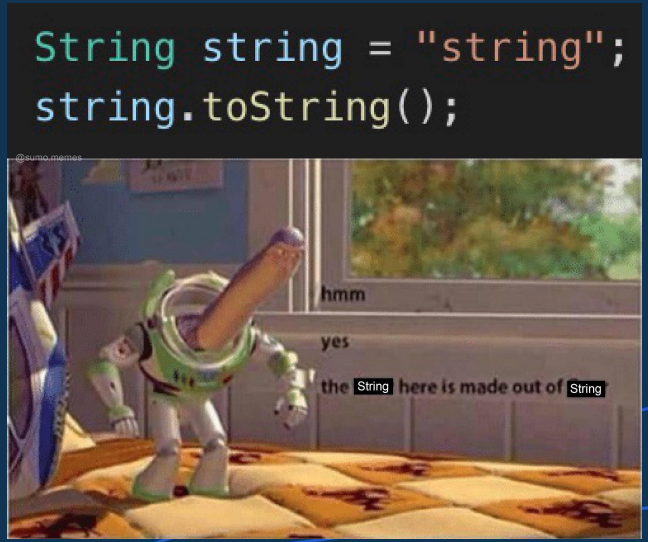
Disadvantages of Strings

- Immutability
- Memory Consumption
- Performance cost
- Encoding and Decoding Overhead
- Security Vulnerabilities



Applications of Strings

- Information Retrieval
- Data analysis
- Network communication
- File handling
- Encoding/Decoding



More applications of Strings

- Bioinformatics
- Search Engines
- Spam Detection

Quiz: Strings

A string is a sequence of characters:

👍True

😂False

Quiz: Strings

Strings are not available as primitive types and in others as composite types:

👍True

😂False

Quiz: Strings

Which of the following is an advantage of strings:

👍Strings are easy to use and manipulate

😂In many programming languages, strings are immutable

Quiz: Strings

It is possible to add two strings together:

👍True

😂False

Character encoding

Character encoding tells computers how to interpret digital data into letters, numbers and symbols.

Character Encoding 101 by Kaðlín Örvardóttir

may display like this:

Character Encoding 101 by Ka□l?n □rvard?ttir

ASCII

(American Standard Code for Information Interchange)

ASCII Table

Dec	Hex	Oct	Char	Dec	Hex	Oct	Char	Dec	Hex	Oct	Char	Dec	Hex	Oct	Char
0	0	0		32	20	40	[space]	64	40	100	@	96	60	140	`
1	1	1		33	21	41	!	65	41	101	A	97	61	141	a
2	2	2		34	22	42	"	66	42	102	B	98	62	142	b
3	3	3		35	23	43	#	67	43	103	C	99	63	143	c
4	4	4		36	24	44	$	68	44	104	D	100	64	144	d
5	5	5		37	25	45	%	69	45	105	E	101	65	145	e
6	6	6		38	26	46	&	70	46	106	F	102	66	146	f
7	7	7		39	27	47	'	71	47	107	G	103	67	147	g
8	8	10		40	28	50	(	72	48	110	H	104	68	150	h
9	9	11		41	29	51	)	73	49	111	I	105	69	151	i
10	A	12		42	2A	52	*	74	4A	112	J	106	6A	152	j
11	B	13		43	2B	53	+	75	4B	113	K	107	6B	153	k
12	C	14		44	2C	54	,	76	4C	114	L	108	6C	154	l
13	D	15		45	2D	55	-	77	4D	115	M	109	6D	155	m
14	E	16		46	2E	56	.	78	4E	116	N	110	6E	156	n
15	F	17		47	2F	57	/	79	4F	117	O	111	6F	157	o
16	10	20		48	30	60	0	80	50	120	P	112	70	160	p
17	11	21		49	31	61	1	81	51	121	Q	113	71	161	q
18	12	22		50	32	62	2	82	52	122	R	114	72	162	r
19	13	23		51	33	63	3	83	53	123	S	115	73	163	s
20	14	24		52	34	64	4	84	54	124	T	116	74	164	t
21	15	25		53	35	65	5	85	55	125	U	117	75	165	u
22	16	26		54	36	66	6	86	56	126	V	118	76	166	v
23	17	27		55	37	67	7	87	57	127	W	119	77	167	w
24	18	30		56	38	70	8	88	58	130	X	120	78	170	x
25	19	31		57	39	71	9	89	59	131	Y	121	79	171	y
26	1A	32		58	3A	72	:	90	5A	132	Z	122	7A	172	z
27	1B	33		59	3B	73	;	91	5B	133	[	123	7B	173	{
28	1C	34		60	3C	74	<	92	5C	134	\	124	7C	174	|
29	1D	35		61	3D	75	=	93	5D	135	]	125	7D	175	}
30	1E	36		62	3E	76	>	94	5E	136	^	126	7E	176	~
31	1F	37		63	3F	77	?	95	5F	137	_	127	7F	177	

Unicode

Unicode is an international character encoding standard that provides a unique number for every character across languages and scripts, making almost all characters accessible across platforms, programs, and devices.

CHARACTER	CODE POINT
A	U+0041
a	U+0061
0	U+0030
9	U+0039
!	U+0021
Ø	U+00D8
ƃ	U+0683
ಚ	U+0C9A
𠜎	U+2070E
😁	U+1F601

Pascal strings

Pascal string is a sequence of characters with optional size specification. It may contain numeric characters, letters, blanks, special characters or a combination of all of them.

```
program exampleStrings;
var
  string1: string[10];  // Character array string
  string2: string;  // String variables
  string3: ShortString // Short string
  string4: pchar; // Null terminated string
  string5: ansistring; // AnsiString
```

Strings in C/C++

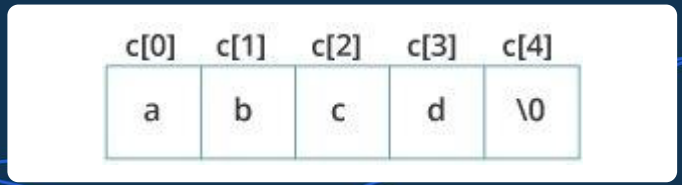
In C/C++, a string is a sequence of characters terminated with a null character *\0*.

```
char myString[] = "abcd";

char myString[10] = "abcd";

char myString[] = {'a', 'b', 'c', 'd', '\0'};

char myString[5] = {'a', 'b', 'c', 'd', '\0'};
```



Strings in C

Common functions

- strlen()
- strcpy()
- strcat()
- strlwr()
- strupr()

```c
// C program to illustrate strings

#include <stdio.h>
#include <string.h>

int main()
{
   char str1[] = "HyperionDev";

   char str2[] = "Software";

   int length = strlen(str1);

   strcpy(str2, str1);

   strcat(str2, str1);

   char str3[] = strlwr(str1);

   char str4[] = strupr(str2);

   return 0;
}
```

Strings in C++

Common functions

- length()
- strcpy()
- strcat()
- tolower()
- toupper()

```
// C++ program to illustrate strings

#include <iostream>
#include <string>
#include <cstring>

int main()
{
   string str1 = "HyperionDev";

   String str2 = "Software";

   int length = str1.length();

   strcpy(str2, str1);

   strcat(str2, str1);

   string str3 = tolower(str1);

   string  str4 = toupper(str2);

   return 0;
}
```

Strings in Java

In java, objects of String are immutable which means a constant and cannot be changed once created.
StringBuilder in Java represents a mutable sequence of characters.

```
// String literal
String myString1 = “abcd”;

// using new keyword
String myString2 = new String("abcd");

// using StringBuilder() constructor
StringBuilder myString3 = new StringBuilder();

// using StringBuilder(CharSequence) constructor
StringBuilder myString4 = new StringBuilder("abcd");

// using StringBuilder(capacity) constructor
StringBuilder myString5 = new StringBuilder(10);

// using StringBuilder(String) constructor
 StringBuilder myString6 = new StringBuilder(myString1.toString());
```

Strings in Java

Some functions / methods

- length()
- valueOf()
- concat()
- toLowerCase()
- toUpperCase()

```
// Java program to illustrate strings

public class Main {
   public static void main(String[] args) {
      String str1 = "HyperionDev";

      String str2 = new String("Software");

      int length = str1.length();

      String str3 = String.valueOf(str1);

      String str4 = str1.concat(str2);

      String str5 = str1.toLowerCase();

      String str6 = str2.toUpperCase();
   }
}
```

Strings in Python

Python strings represents a sequence of characters that are immutable.

```
# Python program to illustrate strings

string_1 = 'abcd'

string_2 = "abcd"

string_3 = '''abcd'''
```

Strings in Python

Some functions / methods

- len()
- index()
- strip()
- lower()
- upper()

```
# Python program to illustrate strings

string_1 = 'HyperionDev'

string_2 = "Software"

string_3 = '''    Web    '''

string_length = len(string_1)

string_index = string_2.index('w')

string_4 = string_3.strip()

string_5 = string_1.lower()

string_5 = string_2.upper()
```

Quiz: Strings

Character encoding tells computers how to interpret digital data into letters, numbers and symbols:

👍True

😂False

Quiz: Strings

ASCII stands for:

👍American Standard Characters for Information Interchange

😂American Standard Code for Information Interchange

Quiz: Strings

Before Unicode, there were hundreds of different character encodings for assigning letters and other characters to a number that could be read by a computer:

👍True

😂False

Quiz: Strings

Strings in C programming language is a sequence of characters terminated with a newline character ‘\n’:

👍True

😂False

Challenge:

Strings

- Reverse words in a given string without using string slicing

- Create program to convert roman numerals to decimal numbers between 1 to 3999

- Create a program to check if a given string is a pangram or not:
  A pangram is a sentence containing every letter in the English alphabet.

  Example:
  “The quick brown fox jumps over the lazy dog”

Q & A

THANK YOU

HyperionDev