

CS250 Intro to CS II

Shereen Khoja

What I think you know

- Variables & data types
- Relational & logical operators
- Decision statements
- Repetition statements including nesting
- Functions
- Files
- Arrays
- Character processing

Review

- Review Reading: pp. 265-284 (files)
pp. 377-425 (arrays)
- Topics to Review
 - Files (Reading & Writing)
 - Arrays (1D & 2D)
 - Character Processing

Files

- What is a stream?
- How do we create a stream for reading from a file?

One-dimensional arrays

- Consider

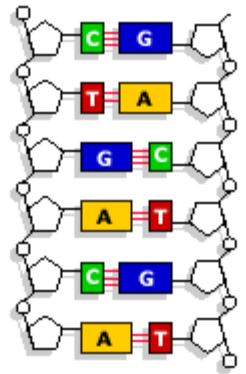
```
int intArray[] = {1, 2, 3, 4, 5};
```

```
double doubleArray[10];
```

1. What are the index values for each array?
2. How many elements does each array have?
3. Arrays consist of *homogeneous* data. What does this mean?

Problem #1

DNA Length



- A DNA string represents the order of nucleobases along one strand of a double-stranded DNA molecule
- The other strand is the ***reverse complement*** of the string
- DNA strings are constructed from the alphabet {A, C, G, T} representing the bases adenine, cytosine, guanine, and thymine

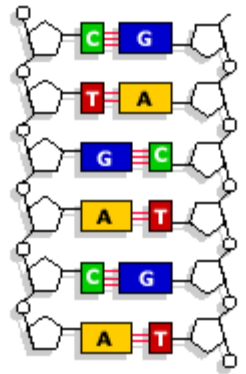
Problem #1

DNA Length

- The DNA string AAGATGCCGT has length 10 nucleobases (or just bases)
- Write a function **getDNALength** that accepts a character array (that ends with null) and returns the length of a DNA string

Problem #2

Reverse Complement



- In DNA strings, the symbols A and T are complements of each other as are C and G
- The reverse complement of a DNA string s is formed by reversing the contents of s and then taking the complement of each symbol
- The DNA string AAAACCCGGT has the reverse complement ACCGGGTTTT

Problem #2

Reverse Complement

- The data file dnastrings.txt contains an unknown number of DNA strings where each string has length at most 1000 bases
- Write a program to output
 - each DNA string
 - the reverse complement of each DNA string
- Let's write
 - well-defined function prototypes
 - each function definition
 - main function for solving the stated problem