

# Strings

Problem Solving Club 2016

# What is the time complexity?

- `char charAt(int index)`
  - Returns the char value at the specified index.
- `charAt` runs in  $O(1)$
- `boolean equals(String otherString)`
  - Returns whether string equals the other one.
- `equals` runs in  $O(n)$
- `boolean contains(CharSequence s)`
  - Returns true if and only if this string contains the specified sequence of char values.
- `contains` runs in  $O(nm)$ . Note: There is an algorithm that runs in  $O(n+m)$  called KMP.

# Regular expressions

- A sequence of characters that define a search pattern, mainly for use in pattern matching with strings.
- `ab*c`
- `cpsc\d+`
- Lots of theory related to regular expressions and regular languages

# Ad-Hoc String Problems

- Some common types of string-related competitive programming problems:
  - Cipher/Encode/Decode
  - Frequency Counting
  - Parsing (Recursive or Non-Recursive)
  - Output formatting
  - Requiring advanced string algorithms
    - KMP: Find a string within another string in  $O(n+m)$
    - Aho-Corasick: Find any number of strings in another string. Used by anti-virus scanners.

# Example problem

- Parentheses Balance:  
<https://uva.onlinejudge.org/external/6/673.pdf>
- Identifying Legal Pascal Real Constants:  
<https://uva.onlinejudge.org/external/3/325.pdf>