Strings

Goldsmiths Computing

Motivation

Most language text is linear, so it makes sense to be able to store text in a linear collection, which we call strings.

Definition

A string is a linear collection specialized to hold characters.
(but what meaning of “character”? Usually code point)

Implementation

For now:

- as dynamic array of code points

  C++ `std::string`

- as vector of code points

  C `char []`

- as *immutable* vector of code points

  Java `java.lang.String`

but beware:

- these data structures might not be optimal for the job
- there are many more exotic implementations and representations out there

Operations

Linear collection operations:

length	return how many characters are in the string
get[i]	return the character at position i in the string
find[c]	is the character c in the string?
position[c]	what position is the character c at?

Mutable collection operations:

push	add a character at the end (C++ only)

String operations:

match[s]	is the string s contained in the string?