BSc (Hons) Computer Science
University of Portsmouth
Third Year

Theoretical Computer Science

M21276
Semester 1

Hugh Baldwin
hugh.baldwin@myport.ac.uk

Contents

Lecture - Induction Lecture

13:00	25/09/24	Janka Chlebikova

Contact Time

- Lectures
  - Delivered by Janka
  - Slides and videos from Covid times available weekly on Moodle
- Tutorials
  - Delivered by either Janka or Dr Paolo Serafino
  - Worksheets available with solutions weekly on Moodle
  - 5 Groups
  - Work through the worksheet solutions
  - Kahoot quiz for revision
  - ‘Tutorial 0’ available as revision of concepts from DMAFP needed for this module

Assessments

- 50%, 90 minute In-Class Test covering Part A - November 20th, 2pm
- 50%, 90 minute Exam covering Part B (Only) - TB1 Assessment Period (January)
- Deferred assessment covers both Part A & B, so will be harder than the two separate exams

There are two past papers (paper-based, but still relevant to the computer-based test) available for each of the assessments. As with DMAFP, the solutions are supplied in the form of a crowd-sourced document that Janka will check over before the exam.

Resources

Lecture slides, videos and tutorial papers are released every week. All of the books on the reading list are available online through the links on Moodle.

Lecture - A1: Introduction to Languages

13:00 03/10/24 Janka Chlebikova

Languages

In this context, a **language** is a set of symbols which can be combined to create a list of acceptable **strings**. There are then also rules which tell us how to combine these strings together, known as **grammars**. The combination of an alphabet, list of valid strings and grammars is known as a language. There will be a formal definition later on.

Definition

An **Alphabet** is a finite, non-empty set of symbols.

For example, in the English language we would define the alphabet, A, as $\mathrm{A} = \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}, \ldots, \mathrm{x}, \mathrm{y}, \mathrm{z}\}$. These symbols can then be combined to create a **string**.

Definition

A **String** is a finite sequence of symbols from the alphabet of a language.

With the alphabet A, we could have strings such as ‘cat’, ‘dog’, ‘$\mathrm{antidisestablishmentarianism}$’, etc *over* the alphabet. A string with no symbols, and therefore a length of zero, is known as the **empty string**, and is denoted by Λ (capital lambda).

Definition

A **Language** over an alphabet (e.g. English over A) is a set of strings – including Λ – made up of symbols from A which are considered ‘valid’. They could be valid as per a set of rules, or could be arbitrary as in most spoken languages.

If we have an alphabet, A, then $\mathrm{A}*$ denotes the infinite set of all strings made up of symbols in A – including Λ.