# Fundamentals of Computer Science for Linguistics

## Dott. Davide Girompini, Dott. Paolo Ruffolo

***COURSE AIMS AND INTENDED LEARNING OUTCOMES***

The overall aim of this course is to provide students with the basic information, terminology, and concepts of computer science and to provide also introductory practical elements on the use of computer for linguistic analysis.

The course will be structured in 2 separate modules that will be delivered one after the other. The first module will be covering the introduction of computer science and information technology including elements of the relevance of information systems in today’s digital value chains.

In the second module an introduction to the GNU/Linux based operating systems Command Line Interface (CLI) will be given. Commonly available Free and Open Source (FOSS) programs will be presented along with some basic technics to combine them to solve simple problems. A short introduction to scripting and programming in C, Perl and Python will be also given.

At the end of the first module, students will be able to understand the computer science basic concepts and components and to comprehend the relevance and the role of information technology in today’s working environments.

At the end of the second module students will know how to use CLI to control their system and to solve simple problems by means of FOSS software, to correctly read software manuals and to interact with repositories. They will also be able to write from scratch very simple programs and scripts on their own and to figure out what a generic script or program actually does.

*COURSE STRUCTURE*

The first module (30 hours, five weeks) will occupy the first 5 weeks of the first semester and will be followed by a week of suspension of lessons where an optional exam session will be held to cover just the first module.

The second module will be delivered on the following 5 weeks (30 hours).

In the normal exam sessions students will bring both modules unless the first module was certified in the optional session on the sixth week.

***COURSE CONTENT***

***COURSE CONTENT***

After a brief historical introduction, the first module will cover the following main subjects:

*Basics*

– Hardware, software, CPU, RAM, storage

– Machine language and assembler

– Operating systems; operational software and application software

– Free and open-source vs proprietary software

*Systems architectures*

– Architecture layers

– Architecture models

– Cloud architectures

*Data and data storing*

– Structured and unstructured data

– Database evolution

– Big data and business intelligence

– Advanced analytics and Artificial Intelligence

– Data integrity, data security, data privacy

*Coding languages and approaches*

– Programming lifecycle: from requirements to solutions

– Informal representations: pseudo-code and flow chart

– Programming paradigms

– Object-oriented programming and micro-services

*Digital value chains*

– Role of information and computing

– Platform business and skill blending

– Role and relevance of linguistic skills for artificial intelligence

The second module will cover the following main subjects:

*GNU/Linux CLI*

– GNU Project and Linux

– Where’s my CLI: availability and installation

– Basic usage tips

– The most important command: man

– GNU Bash

– File system and generic commands

*Text and Data manipulation*

– Regular Expressions Basics

– *sed*, *grep*, *awk*

– *nano* and *vi*

*Basic Bash scripting*

– Variables

– Control Flow

– Debugging

*De-Facto Standard FOSS programs*

– *ssh*, *scp*: using remote machines

– *git*: interacting with repositories

*Programming Languages*

– Interpretation And Translation

– GNU Compiler Collection (*gcc*)

– *Perl*

– *Python*

– *C*

***READING LIST***

All the readings reported below are not mandatory for the exam. Most of them are available in several formats including printed version and, for some of them, also the Italian version. Additional optional consultation material will be provided during the course.

J. Valacich-C. Schneider, Information Systems Today, Managing in the Digital World – ‎Pearson, eight edition – ISBN: ‎978-1292215976

B. Kernighan, Understanding the Digital World: What You Need to Know about Computers, the Internet, Privacy, and Security, Second Edition – ‎Princeton Univ Pr; second edition (March 30, 2021) - ISBN: ‎978-0691219103

AA.VV., Introduction to the Command Line ISBN: 978-1-882114-04-7 [*https://static.fsf.org/nosvn/gnu-press/source-files/books/command-line.pdf*](https://static.fsf.org/nosvn/gnu-press/source-files/books/command-line.pdf)

B.W. Kernighan-D.M. Ritchie, C Programming Languag – Pearson; second edition (March 22, 1988) - ISBN:‎ 978-0131103627

Al Aho-J. Ullman, Foundations of Computer Science – Computer Science Press; first edition (January 1, 1992) - ISBN: 978-0716782339 - [*http://infolab.stanford.edu/~ullman/focs.html*](http://infolab.stanford.edu/~ullman/focs.html)

GNU coreutils [*https://www.gnu.org/software/coreutils/manual/*](https://www.gnu.org/software/coreutils/manual/)

GNU Bash manual, [*https://www.gnu.org/software/bash/manual/*](https://www.gnu.org/software/bash/manual/)

Perl Documentation, [*https://www.perl.org/docs.html*](https://www.perl.org/docs.html)

Python Documentation, [*https://docs.python.org/3/tutorial/*](https://docs.python.org/3/tutorial/)

Git documentation, [*https://git-scm.com/docs*](https://git-scm.com/docs)

Openssh documentation, [*http://www.openssh.com/manual.html*](http://www.openssh.com/manual.html)

***TEACHING METHOD***

Lectures with exercises.

***ASSESSMENT METHOD AND CRITERIA***

For the first module, both in the intermediate and regular exam sessions:

There will be 5 written questions with multiple choices and one final oral discussion where the answers will be reviewed, and additional questions will be asked.

For the second module:

By means of a written test students will be asked to write a simple script/program to solve a simple problem and/or to describe what is the expected output/effects of a proposed program/script. Test will be discussed in the following oral exam.

The exams for the two modules can be taken in any order and the overall result of the course’s exam will be the arithmetic average of the two modules’ results.

***NOTES AND PREREQUISITES***

Given the introductory nature of the course, no prior specific expertise in computer science is required, apart from basic competence (e.g. using a web browser).

*Place and time of consultation hours*

On appointment, by sending an e-mail to:

davide.girompini1@unicatt.it for the first module

paolo.ruffolo@unicatt.it for the second module

The consultation can be also separate for each single module

CIRCSE Research Center, Franciscanum building, second floor, room n. 218.