**Introduction to Scientific Computing**

**7-11 October 2019**

The modern world is ever more dependent on computer science and information technology advances. Especially in life science in recent years, explosive amounts of biological information have been obtained and deposited in various databases. The predominant source of this data is “high-throughput” experimentation, involving simultaneous execution of hundreds or thousands of experiments. A comprehensive understanding of biological phenomena can be achieved only through the integration of all available biological information and different data analysis tools and applications. It becomes crucial for modern scientists to acquire skills that will enable to deal with this data explosion. It is not surprising that scientists with strong computer skills fare better in their position and research. This course aims to provide learners with the basic notions underlying computer science. The focus of this course in on computer architecture, scientific computing and basic programming.

**Aim of the course:**

To provide participants with basic knowledge on computer architecture and available resources, programming in MATLAB, and good IT practices.

**By the end of the course, the participants should be able to:**

* Describe personal computer & cluster architecture and their usage.
* Describe the main operational systems & programming languages
* Write basic scripts and functions in MATLAB
* Describe some principles of good programming practice
* Name the available IT resources and people in charge at the GIGA/ULiège

**Target group:** PhD candidates. Course is limited to 15 participants

**Prerequisites:** Working knowledge of English; personal laptop; preferably with MATLAB pre-installed, or it will be installed on Day 1.

**Duration of the course and workload:** 5 days; 38 hours

**Organization of the course:**

Except for Day 1, each morning is spent discussing the material studied and exercises solved the previous afternoon. A one hour of lecture on a specific topic is also scheduled. In the afternoon, attendees will be expected to follow MOOC videos (MATLAB and Octave for Beginners, edX platform) and read a few chapters from a MATLAB book. Then they will try and solve some problems by writing MATLAB code in pairs.

**Location:** B34, +5 GIGA

**Educators:**

Christophe Phillips [CP], Mohamed Ali Bahri [MB], Gregory Hammad [GH], Pierre-François Pirlet [PFP] (GIGA- CRC in vivo imaging and GIGA-In silico Medicine units)

**Course program:**

**Day 1**

**9:00 - 12:30 : Basics of computer science (lectures)**

1. Introduction to computer science: historical perspective, computer structure, operating systems & languages [CP]
2. Data representation & storage: bits/bytes, data format, signal discretization, compression [MB]
3. Mass storage & ULiège IT resources

**12:30 - 14:00** **Lunch break**

**14:00 - 15:00 Setting up MATLAB programming course**

Online resources presentation: theory and practicals [MB, CP]

Pairing for the practicals

**15:00 - 17:00 Introduction to MATLAB programming & exercises, Part 1**

Chapters 1 to 4 & video (1.1-3; 2.1-5). Exercises

**Day 2**

**9:00 - 11:00 Introduction to MATLAB programming** [MB, CP]

Q&A about previous day’s material. Solution to the exercises

**11:30 - 12:30 Introduction to algorithmics (lecture)**  [MB]

 Different types of algorithms, computer logic

**12:30 - 14:00** **Lunch break**

**14:00 - 17:00 Introduction to MATLAB programming & exercises, Part 2**

Chapters 5 to 8 & video (3.1; 4.1-3). Exercises

**Day 3**

**9:00 - 11:00 Introduction to MATLAB programming** [MB, CP]

Q&A about previous day’s material. Solution to the exercises

**11:30 - 12:30 Good practices in scientific computing (lecture)** [CP]

 Code modularity, validation, optimization, documentation

**12:30 - 14:00** **Lunch break**

**14:00 - 17:00 Introduction to MATLAB programming & exercises, Part 3**

Chapters 9 to 11 & video (3.2-4). Exercises

**Day 4**

**9:00 - 11:00 Introduction to MATLAB programming** [MB, CP]

Q&A about previous day’s material. Solution to the exercises

**11:30 - 12:30 Versioning & Git (lecture)** [CP/GH]

 Problems, solution & tools

**12:30 - 14:00** **Lunch break**

**14:00 - 17:00 Introduction to MATLAB programming & exercises, Part 4**

Chapters 12 to 14 & video (5.1, 5.2)**.** Exercises

**Day 5**

**9:00 - 11:00 Introduction to MATLAB programming** [MB, CP]

 Q&A about previous day’s material. Solution to the exercises

**11:30 - 12:30 Working with clusters (lecture)** [GH]

 Working on cluster, parallel processing, data management, resources.

**12:30 - 14:00** **Lunch break**

**14:00 - 15:00 Working with personal data ? Think**[**GDPR**](https://en.wikipedia.org/wiki/General_Data_Protection_Regulation)**!** [PFP]

**15:00 - 17:00 Working with Big Data**

A series of talks by the GIGA researches on working with big data. The focus of the talks:

* The application field and scale of the problem
* The computational resources needed
* The data flow
* Key elements in the processing

**Support resources:**

* MOOC: [MATLAB and Octave for Beginners](https://www.edx.org/course/matlab-octave-beginners-epflx-matlabeoctavebeginnersx) This will be the main MOOC about MATLAB programming. The whole pdf book for the course (chapters to read and exercices) will be made available or it can be purchased on [Amazon](https://www.elsevier.com/books/matlab/attaway/978-0-12-804525-1)

Additional resources:

* MOOC: [IT support: Hardware essentials](https://www.edx.org/course/it-support-hardware-essentials-0)
* MOOC: [Introduction to Linux](https://www.edx.org/course/introduction-linux-linuxfoundationx-lfs101x-1)
* MOOC: [Data Science: Productivity tools](https://www.edx.org/course/data-science-productivity-tools-harvardx-ph125-5x)