Course guide
270540 - TEB - Techniques and Tools for Bioinformatics

Last modified: 03/02/2025

Unit in charge: Barcelona School of Informatics
Teaching unit: 723 - CS - Department of Computer Science.

Degree: MASTER'S DEGREE IN INFORMATICS ENGINEERING (Syllabus 2012). (Optional subject).
MASTER'S DEGREE IN INNOVATION AND RESEARCH IN INFORMATICS (Syllabus 2012). (Optional subject).
MASTER'S DEGREE IN DATA SCIENCE (Syllabus 2021). (Optional subject).

Academic year: 2024 **ECTS Credits:** 3.0 **Languages:** English

LECTURER

Coordinating lecturer: GABRIEL ALEJANDRO VALIENTE FERUGLIO

Others:

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CDG1. Capability to integrate technologies, applications, services and systems of Informatics Engineering, in general and in broader and multicisciplinary contexts.
CTE7. Capability to understand and to apply advanced knowledge of high performance computing and numerical or computational methods to engineering problems.
CTE9. Capability to apply mathematical, statistical and artificial intelligence methods to model, design and develop applications, services, intelligent systems and knowledge-based systems.

Generical:
CG4. Capacity for mathematical modeling, calculation and simulation in technology and engineering companies centers, particularly in research, development and innovation tasks in all areas related to Informatics Engineering.

Transversal:
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.

Basic:
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

TEACHING METHODOLOGY

In the theoretical sessions, the lecturer will introduce algorithms and data structures, combined with examples and problem-solving. In the problem-solving sessions, students will work on their own solving problems, under supervision and assistance of the lecturer.

LEARNING OBJECTIVES OF THE SUBJECT

1.Understand pattern matching algorithms and data structures and their implementation in a modern programming language, and apply them to solve practical problems in bioinformatics.

STUDY LOAD

Type	Hours	Percentage
Hours large group	13,5	18.00
Hours small group	13,5	18.00
Self study	48,0	64.00

Total learning time: 75 h

CONTENTS

Efficient search algorithms and data structures

Description:
The most efficient algorithms for searching for patterns in texts based on the alphabet and the length and number of patterns will be shown. Data structures that are very useful for comparing genomes such as suffix trees, suffix arrays and the Burrows-Wheeler transform, will also be explained.

Sequence alignment

Description:
Dynamic programming will be explained and we will discuss its application to compute the edit distance between two words, to the approximate search of a word in a text and to find the best alignment between two sequences. We will also study how to generalize the alignment to multiple sequences.

Data base searching: BLAST

Description:
The computational and statistical foundations of the BLAST algorithm and its use for approximate searches in databases will be introduced.

Alignment of degenerate strings

Description:
The generalization to pangenomes of pattern search problems will be explained and the extensions of the Boyes-Moore algorithm and the Burrows-Wheeler transform for the alignment of degenerate sequences will be introduced.

ACTIVITIES

Efficient search algorithms and data structures

Specific objectives:
1

Related competencies :
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CDG1. Capability to integrate technologies, applications, services and systems of Informatics Engineering, in general and in broader and multicisciplinary contexts.
CTE7. Capability to understand and to apply advanced knowledge of high performance computing and numerical or computational methods to engineering problems.
CTE9. Capability to apply mathematical, statistical and artificial intelligence methods to model, design and develop applications, services, intelligent systems and knowledge-based systems.
CG4. Capacity for mathematical modeling, calculation and simulation in technology and engineering companies centers, particularly in research, development and innovation tasks in all areas related to Informatics Engineering.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.

Full-or-part-time: 20h
Self study: 12h
Theory classes: 4h
Practical classes: 2h
Laboratory classes: 2h

Sequence alignment

Specific objectives:
1

Related competencies :
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CDG1. Capability to integrate technologies, applications, services and systems of Informatics Engineering, in general and in broader and multicisciplinary contexts.
CTE7. Capability to understand and to apply advanced knowledge of high performance computing and numerical or computational methods to engineering problems.
CTE9. Capability to apply mathematical, statistical and artificial intelligence methods to model, design and develop applications, services, intelligent systems and knowledge-based systems.
CG4. Capacity for mathematical modeling, calculation and simulation in technology and engineering companies centers, particularly in research, development and innovation tasks in all areas related to Informatics Engineering.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.

Full-or-part-time: 20h
Self study: 12h
Theory classes: 4h
Practical classes: 2h
Laboratory classes: 2h

Database searches: BLAST

Specific objectives:
1

Related competencies :
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CDG1. Capability to integrate technologies, applications, services and systems of Informatics Engineering, in general and in broader and multicisciplinary contexts.
CTE7. Capability to understand and to apply advanced knowledge of high performance computing and numerical or computational methods to engineering problems.
CTE9. Capability to apply mathematical, statistical and artificial intelligence methods to model, design and develop applications, services, intelligent systems and knowledge-based systems.
CG4. Capacity for mathematical modeling, calculation and simulation in technology and engineering companies centers, particularly in research, development and innovation tasks in all areas related to Informatics Engineering.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.

Full-or-part-time: 10h
Self study: 6h
Theory classes: 2h
Practical classes: 1h
Laboratory classes: 1h

Alignment of degenerate strings

Specific objectives:
1

Related competencies :
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CDG1. Capability to integrate technologies, applications, services and systems of Informatics Engineering, in general and in broader and multicisciplinary contexts.
CTE7. Capability to understand and to apply advanced knowledge of high performance computing and numerical or computational methods to engineering problems.
CTE9. Capability to apply mathematical, statistical and artificial intelligence methods to model, design and develop applications, services, intelligent systems and knowledge-based systems.
CG4. Capacity for mathematical modeling, calculation and simulation in technology and engineering companies centers, particularly in research, development and innovation tasks in all areas related to Informatics Engineering.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.

Full-or-part-time: 10h
Self study: 6h
Theory classes: 2h
Practical classes: 1h
Laboratory classes: 1h

Presentation of advanced algorithms and data structures

Specific objectives:
1

Related competencies :
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CDG1. Capability to integrate technologies, applications, services and systems of Informatics Engineering, in general and in broader and multicisciplinary contexts.
CTE7. Capability to understand and to apply advanced knowledge of high performance computing and numerical or computational methods to engineering problems.
CTE9. Capability to apply mathematical, statistical and artificial intelligence methods to model, design and develop applications, services, intelligent systems and knowledge-based systems.
CG4. Capacity for mathematical modeling, calculation and simulation in technology and engineering companies centers, particularly in research, development and innovation tasks in all areas related to Informatics Engineering.
CTR4. INFORMATION LITERACY: Capability to manage the acquisition, structuring, analysis and visualization of data and information in the area of informatics engineering, and critically assess the results of this effort.

Full-or-part-time: 15h
Self study: 15h

GRADING SYSTEM

There will be a (mid-term) final exam, in which the students will explain advanced algorithms and data structures from the research literature. The final grade will be just the final exam grade.

BIBLIOGRAPHY

Basic:
- Mäkinen, Veli; Belazzougui, Djamal; Cunial, Fabio; Tomescu, Alexandru I. Genome-scale algorithm design : bioinformatics in the era of high-throughput sequencing. 2nd ed. Cambridge: Cambridge University Press, 2023. ISBN 9781009341233.

Complementary:
- Mount, David W. Bioinformatics : sequence and genome analysis. 2nd ed. New York: Cold Spring Harbor Laboratory Press, cop. 2004. ISBN 0879696877.