GHENT UNIVERSITY

Course Specifications

Valid as from the academic year 2024-2025

Bioinformatics 1 (0000178)

Course size *(nominal values; actual values may depend on programme)*

Credits 5.0 **Study time 150 h**

Course offerings and teaching methods in academic year 2024-2025

A (semester 1)	English	Incheon	lecture
			seminar

Lecturers in academic year 2024-2025

De Neve, Wesley	KR01	lecturer-in-charge
Özbulak, Utku	KR01	co-lecturer

Offered in the following programmes in 2024-2025	**crdts**	**offering**
Bachelor of Science in Molecular Biotechnology	5	A

Teaching languages

English

Keywords

Algorithmic design, Algorithms, Bioinformatics, Creative problem solving, Complexity analysis, Datastructures, Programming, Python

Position of the course

The field of bioinformatics was born after biologists discovered how to sequence (digitize) DNA, raising the need for mathematical and computational techniques to decipher the language of DNA, RNA, and proteins. As a result, bioinformatics has become an important part of modern biology, often facilitating new insights and new (data-driven) approaches, driving forward biological research and development efforts.

Primarily taking an algorithmic point-of-view, this course aims at introducing students to the design, implementation, and analysis of standard problem-solving techniques in the field of bioinformatics, including exhaustive search algorithms, recursive algorithms, divide-and-conquer algorithms, greedy algorithms, graph algorithms, dynamic programming algorithms, randomized algorithms, and machine-learning algorithms. These algorithms and related datastructures (e.g., lists, tuples, sets, dictionaries, graphs, hash tables, and trees) are studied in the context of problems like pattern finding, genome rearrangements, genome assembly, sequence alignment and comparison, regulatory motif finding, and genome annotation.

Contents

Based on Capita Selecta from the course handbook used, this course deals with the following computational topics, making use of pseudocode and the Python programming language:

- algorithms and complexity, including asymptotic notation (Chapter 1, Appendix);
- exhaustive search algorithms and pruning (Chapter 1, Chapter 4);
- recursive and divide-and-conquer algorithms (Chapter 1, Chapter 5, Appendix);
- greedy algorithms (Chapter 6);
- graph algorithms (Chapter 3);
- dynamic programming algorithms (Chapter 5);
- randomized and genetic algorithms (Chapter 2); and
- machine learning algorithms (slides).

The above algorithms are studied in the context of the following problems:

- pattern finding in biological sequences;
- genome rearrangements;
- genome assembly;
- sequence alignment and comparison;
- regulatory motif finding; and
- genome annotation.

The theory is brought into practice through different series of mostly computer-based exercises.

Initial competences

An understanding of basic probability theory and the behaviour of common functions studied in calculus (e.g., linear functions, logarithmic functions, exponential functions).

Prior experience in at least one high-level programming language (e.g., Python).

Informatics (0000096)

Final competences

1 Understand the biological motivation of common bioinformatics problems, as well as their mathematical and computational abstraction.
2 Be familiar with the standard algorithms and datastructures used in bioinformatics.
3 Design and implement algorithms for new bioinformatics problems using standard algorithmic techniques.
4 Compare the effectiveness and time/memory efficiency of different algorithmic techniques in solving different bioinformatics problems.

Conditions for credit contract

Access to this course unit via a credit contract is determined after successful competences assessment

Conditions for exam contract

This course unit cannot be taken via an exam contract

Teaching methods

Seminar, Lecture

Study material

Type: Handbook

Name: Bioinformatics Algorithms: An Active Learning Approach
Indicative price: € 80
Optional: yes
Author : Phillip Compeau & Pavel Pevzner
ISBN : 978-0-99037-463-3

Type: Slides

Name: Bioinformatics 1
Indicative price: Free or paid by faculty
Optional: no
Language : English
Number of Slides : 550
Available on Ufora : Yes
Available through Student Association : No

Type: Laptop

Name: Laptop
Indicative price: € 1,000
Optional: no
Usability and Lifetime within the Course Unit : intensive
Usability and Lifetime within the Study Programme : regularly
Usability and Lifetime after the Study Programme : regularly

Type: Software

Name: Dodona, Online Python Tutor, PyCharm, Linux shell
Indicative price: Free or paid by faculty
Optional: no

Online Available : Yes

References

Mark Lutz (2009). Learning Python: Powerful Object-Oriented Programming. Fourth Edition. O'Reilly Media, ISBN-13: 978-0596158064.

Mark Pilgrim (2009). Dive into Python. CreateSpace, ISBN-13: 978-1441413024. Free download @ http://diveintopython.org.

Hans Peter Langtangen (2009). A Primer on Scientific Programming with Python. Springer, ISBN-13: 978-3642024740.

Tony Gaddis (2009). Starting Out with Python. Pearson Education - Addison Wesley, ISBN-13: 978-0321549419.

Michael H. Goldwasser (2007). Object-Oriented Programming in Python. Prentice Hall, ISBN-13: 978-0136150312.

Jason Kinser (2008). Python for Bioinformatics. Jones & Bartlett Publishers, ISBN-13: 978-0763751869.

Sebastian Bassi (2009). Python for Bioinformatics. Chapman & Hall, ISBN-13: 978-1584889298.

Mark G. Sobell (2012). A Practical Guide to Linux: Commands, Editors, and Shell Programming. Fourth Edition. Prentice Hall, ISBN-13: 978-0134774602.

William F. Punch, Richard Enbody (2017). The Practice of Computing using Python. Third Edition. Addison Wesley, ISBN-13: ISBN-13: 978-0133085044.

Steven Haddock and Casey Dunn (2010). Practical Computing for Biologists. First Edition. Sinauer Associates, Inc, ISBN-13: 978-0878933914.

Ashley Shade, Tracy K. Teal (2015). Computing Workflows for Biologists: A Roadmap. PLOS Biology.

Pavel A. Pevzner (2004). Educating Biologists in the 21st Century: Bioinformatics Scientists versus Bioinformatics Technicians. Bioinformatics, Vol. 20, No. 14, pages 2159–2161.

Alejandra J. Magana, Manaz Taleyarkhan, Daniela Rivera Alvarado, Michael Kane, John Springer, and Kari Clase (2014). A Survey of Scholarly Literature Describing the Field of Bioinformatics Education and Bioinformatics Educational Research. CBE—Life Sciences Education, Vol. 13, pages 607–623.

Phillip Compeau and Pavel Pevzner (2014). Bioinformatics Algorithms: An Active Learning Approach. Active Learning Publishers, ISBN-13: 978-0990374602.

Neil C. Jones and Pavel A. Pevzner. An Introduction to Bioinformatics Algorithms. The MIT Press, ISBN-13: 978-0262101066.

Istvan Albert (2019). The Biostar Handbook: A Beginner's Guide to Bioinformatics. Self-published. Second Edition.

Course content-related study coaching

The theory lectures touch upon the biological motivation of common bioinformatics problems, as well as their mathematical and computational abstraction.

During the supervised hands-on sessions, students themselves learn how to tackle computational challenges in the field of bioinformatics by solving a series of mandatory exercises. These exercises aim at bringing the theory into practice.

Example examinations of previous years are made available on Ufora near the end of the teaching activities.

Information about the calculation of the different evaluation marks is communicated during the theory lectures at the beginning and near the end of the first-term teaching activities.

Announcements on Ufora are used for counselling, giving feedback, and providing background information.

Through individual appointments scheduled via email, the lecturers and the teaching assistants are available for answering questions about the course in general (grading, examination), the theory, and the exercises.

Assessment moments

end-of-term and continuous assessment

Examination methods in case of periodic assessment during the first examination period

Skills test, Written assessment with multiple-choice questions, Written assessment with open-ended questions, Written assessment

Examination methods in case of periodic assessment during the second examination period

Skills test, Written assessment with multiple-choice questions, Written assessment with open-ended questions

Examination methods in case of permanent assessment

Skills test, Assignment

Possibilities of retake in case of permanent assessment

examination during the second examination period is not possible

Calculation of the examination mark

For the **first-term examination period**, the continuous assessment (hands-on sessions; non-periodic evaluation; permanent evaluation) accounts for 25% of the final examination mark and the end-of-term assessment (examination; periodic evaluation) accounts for 75% of the final examination mark. To qualify for passing, both the mark of the continuous assessment and the mark of the end-of-term assessment should be higher than or equal to 8/20 (40%). If that is not the case, the final examination mark is the least of the two obtained marks. Note that the end-of-term assessment consists of a written closed-book part, accounting for 50% of the mark, and an open-book part, accounting for 50% of the mark.

The continuous assessment cannot be retaken during the **resit examination period**. Therefore, the examination mark for the resit examination period is calculated twice. For the first calculation, the mark of the continuous assessment, as obtained for the first-term examination period, accounts for 25% of the examination mark and the mark of the resit assessment, as obtained during the resit examination period, accounts for the remaining 75% of the examination mark. For the second calculation, the examination mark is equal to the mark of the resit assessment, as obtained during the resit examination period (that is, the mark of the continuous assessment, as obtained for the first-term examination period, is ignored). The final examination mark for the resit examination period is then equal to the maximum of the above two calculations. Note that the resit assessment consists of a written closed-book part, accounting for 50% of the mark, and an open-book part, also accounting for 50% of the mark.

Students who eschew period aligned and/or non-period aligned evaluations for this course unit may be failed by the examiner.