Universities of Minho, Aveiro and Porto

Doctoral Program in Computer Science

"Algorithm Design and Analysis"

April 2009

A. Program

1. Purpose and Justification

This document proposes and describes a **Unit Course on Theory and Fundamentals** for the MAP-I Doctoral Program. This Unit Course is entitled **"Algorithm Design and Analysis"** and will address three major and inter-related themes within the Algorithms and Complexity area, namely, fundamental algorithm design strategies, algorithm efficiency and complexity analysis, and algorithms for Bioinformatics.

This course will provide fundamental knowledge to MAP-I students, who will be able to apply different design strategies and develop appropriate algorithms for different application problems.

The proposed course is aimed at algorithm design strategies and efficiency analysis methodologies, and strives to strike a balance between the required fundamental knowledge and the presentation of different case studies, with emphasis in the Bioinformatics area. The theoretical foundations will be introduced, followed by state-of-the-art approaches to solve practical problems. Emphasis will be put on discussing global design strategies and possible applications. In fact, it is a major goal to stimulate students to explore other application problems and alternative problem solving techniques.

Examples of graduate courses similar to this one can be found in several universities, such as:

- MIT
    - 6.854J — Advanced Algorithms
- CMU
    - 15-750 — Algorithms Core
    - 15-853 — Algorithms in the "Real World"

While the main topics and some of the case problems/studies of the proposed course are the same, the above courses have a larger scope, addressing more advanced complexity analysis methods and encompassing various application areas, and also deal with particular advanced data structures.

The smaller scope of the proposed course is a deliberate option, to accommodate students with less CS background and to allow more emphasis in the Bioinformatics area.

The decision to structure the course around algorithm design strategies (e.g., divide-and-conquer, greedy algorithms, etc.) and not problem types (e.g., sorting, shortest paths, etc.) seems also to be favorable to interested students, by presenting them early with various case problems that can be addressed by different algorithmic strategies along the course.

2. Goals

Since students enrolling in the MAP-I program might have different backgrounds with various levels of knowledge in the fundamental Computer Science area of Algorithms and Complexity, one of the main goals of the proposed course is to help eliminate such differences. Therefore, the proposed course is particularly important for students without an appropriate CS background, and should be attractive to them.

Additionally, and given the theme of the 2008/2009 edition of the doctoral program, the course will focus on particular problems and algorithms in the Bioinformatics area, which will be used to provide students with a more applied perspective. And constitute a first step towards other related courses in the program.

3. Learning Outcomes

Upon successful conclusion of this unit course students will be able to:

- identify, for a certain problem, the more adequate algorithm design strategies;
- design, implement and test an appropriate algorithm for different application problems;
- analyze the efficiency of different algorithms for the same problem and classify them according to their complexity;
- apply the main strategies and algorithms used in particular Bioinformatics problems.

4. Detailed Program

1. **Introduction**
    a. Algorithm Design Strategies
    b. Fundamentals of Algorithm Complexity Analysis
    c. Introduction to Bioinformatics

2. **Algorithm Design**
    a. Brute-Force and Exhaustive Search
    b. Divide-and-Conquer
    c. Decrease-and-Conquer
    d. Dynamic Programming
    e. Greedy Algorithms
    f. Advanced Algorithm Design Techniques
        i. Backtracking and Branch-and-Bound
        ii. Computing Approximate Solutions

3. **Algorithms for Bioinformatics Problems**
    a. String Matching
    b. Suffix Trees and Suffix Arrays
    c. Sequence Alignment
    d. Additional problems / applications

The lectures will provide, for every main course topic, the presentation of the fundamental concepts, some application examples or case studies, as well as suggestions for further work.

The emphasis is on the fundamental concepts, the algorithm design strategies and their applicability to problems of different kinds.

5. Teaching Methodology

The course will be made up of lectures that will be complemented by tutorial meetings between students and professors/researchers, mainly for advice related to the various assignments. The material used by the professors to lecture and support their classes (slides, videos, notes, etc.) will be made available to the students on the course site.

Practical assignments will be defined at the end of each module and may require either development of simple applications, or writing a report about a given theme with associated bibliographic research. Such reports, with a format close to scientific papers, will also serve for their assessment, and encompass state-of-the-art reports, position papers, or discussion of publications. It is expected that students make serious research in this context, using the internationally accepted scientific data bases. Dedicated course workshops can also be organized for presentation, in forum environment, of the work done by the students.

Seminal and fundamental papers will be suggested to students as important reading material and some of them might be presented in class, and discussed in tutorial orientation meetings.

Application areas/problems will be presented and discussed. Besides constituting examples of algorithms usage and allowing the sedimentation of the acquired knowledge as a whole, this is intended as an opportunity for students to exercise their critical abilities and design/propose new ideas and applications.

6. Assessment

Student assessment will be achieved in two main components: project assignments and report writing, and a final examination, each one with a weight of 50%.

At the end of each main topic, a small project assignment will be given. Depending on the topic, the project can be in the design and test of an algorithm or a written report on a particular technique, algorithm, application area, etc.

The report takes the form of a survey and will be oriented for a deeper discussion of several solutions of the elected problem.

Students that fail to perform on any of the above assignment/tasks shall not be considered for final assessment.

7. References

S. Baase and A. Van Gelder. *Computer Algorithms - Introduction to Design and Analysis* - 3rd. Ed. Addison-Wesley, 2000.

H.-J. Böckenhauer and D. Bongartz. *Algorithmic Aspects of Bioinformatics*. Springer, 2007.

T. H. Cormen et al. *Introduction to Algorithms* - 2nd Ed. MIT Press, 2001.

D. Gusfield. *Algorithms on Strings, Trees and Sequences*. Cambridge University Press, 1997.

J. Kleinberg and É. Tardos. *Algorithm Design*. Addison-Wesley, 2005.

A. Levitin. *Introduction to the Design and Analysis of Algorithms*. Addison-Wesley, 2003.

K. Mehlhorn and P. Sanders. *Algorithms and Data Structures — The Basic Toolbox*. Springer

A. Polanski and M. Kimmel. *Bioinformatics*. Springer, 2007.