The MSc in Bioinformatics Program at NTU

Professor Bertil Schmidt and Professor Keong-Kwoh Chee

Background
Bioinformatics deals with the methods for storing, retrieving, and analyzing biological data such as nucleic acid (DNA/RNA) and protein sequences, structures, functions, and pathways, as well as genetic interactions. The demand for bioinformatics is driven by the exponential growth rate of genomic information (e.g. sequence data and microarray data) in recent years. As a consequence, there are currently not enough researchers adequately trained to analyze, manage, and mine these data. Furthermore, there is a need to harness this information for medical diagnostics, therapeutic uses, and other industrial applications. Modern bioinformatics is not simply a subset of biology or computer science: it emphasizes both the domain data and the methods. Hence, it requires specific training beyond either biology or computer science alone.

Teaching Outcome
The ultimate goal of bioinformatics is to uncover the wealth of biological information hidden in the mass of sequence, structure, literature, and other biological data. It therefore aims to obtain a clearer insight into the fundamental biology of organisms, create different levels of abstraction for ease of use and understanding, and eventually use this information to save lives.

There are at least three levels of teaching outcomes for a bioinformatics course program:

- Tool user: using existing programs and methods to answer biologically interesting questions.
- Installer and maintainer: setting up derived databases, translating biologists’ questions into ones that software programs can answer, maintaining software tools, and updating databases regularly.
- Tool builder: creating new or improving existing software tools and methods for analyzing and organizing biological data.

None of these approaches are uniquely “right”. Therefore, there must be a balanced curriculum experience through all levels — from tool users to tool builders. The educational model for bioinformatics is no different from other science and engineering subjects. Although only relatively few professionals are needed to advance the field, all of them must be able to understand and appreciate the tools for profound interpretation of the reported results.

Course Overview
The Master of Science (MSc) program at Nanyang Technological University (NTU) has been designed to address this need. This course offers a two-year part-time training program leading to an MSc. The course is designed for students who have relevant scientific and technical background to upgrade themselves and stay relevant to the fourth pillar of the Singaporean economy: biomedical sciences. The curriculum provides them
with a skill set for the creation of excellent, well-validated methods to solve problems in the domain of bioinformatics and related fields that will emerge as a consequence of Singapore’s drive in biomedical sciences research. The curriculum covers biology, significant coursework in algorithms, biostatistics and probability, data analysis and data mining, as well as a substantial number of computer-related skill and information technology seminars.

**Course Structure**

The part-time course has a normal duration of two years. During this time, students must complete 30 academic units (AUs) to graduate from the course. Three-hour lectures are normally conducted in the evenings. Laboratory or workshop sessions and projects may also be conducted in the weekday evenings or on Saturdays during the day. In order to provide more flexibility to candidates, particularly as the majority of them are working full-time, the following options are available:

1. Coursework and dissertation: Students are required to take eight subjects (24 AUs) consisting of six core subjects and two elective subjects, and a substantial project (6 AUs).
2. Coursework only: Students are required to take a total of ten subjects (30 AUs) consisting of six core subjects and four elective subjects, one of which must be the Directed Reading module (3 AUs).

**Core Subjects**

The program consists of six core subjects, which are described in more detail in Table 1. After the two mandatory introductory courses (Introductory Biology and Introductory Bioinformatics), the students are expected to be proficient in using most bioinformatics tools. After taking all six core subjects, the students are expected to be proficient in implementing, improving, and creating new software tools and methods for analyzing and organizing data.

<table>
<thead>
<tr>
<th>Core Subject Title</th>
<th>Teaching Goals</th>
<th>Recommended Semester</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introductory Biology</td>
<td>Create a framework of essential biological knowledge</td>
<td>Year 1, Sem 1</td>
</tr>
<tr>
<td>Introductory Bioinformatics</td>
<td>Learn about basic bioinformatics concepts, databases, tools, and applications</td>
<td>Year 1, Sem 1</td>
</tr>
<tr>
<td>Computational Biology</td>
<td>Study and apply modern machine learning approaches to computational biology problems</td>
<td>Year 1, Sem 2</td>
</tr>
<tr>
<td>Advanced Biology</td>
<td>Develop a deeper understanding of DNA and gene dynamics as well as biochemistry</td>
<td>Year 1, Sem 2</td>
</tr>
<tr>
<td>Biostatistics</td>
<td>Plan observational studies and mathematically model biological phenomena</td>
<td>Year 2, Sem 1</td>
</tr>
<tr>
<td>Algorithms for Bioinformatics</td>
<td>Become proficient in key algorithmic techniques commonly used in bioinformatics and computational biology</td>
<td>Year 2, Sem 1</td>
</tr>
</tbody>
</table>
Elective Subjects
In addition to the core subjects, the students can choose any of the following electives. The electives offer them the opportunity to specialize in their areas of interest.

- High Performance Computing for Bioinformatics
- Biological Systems Modeling
- Methods and Tools of Proteomics
- Medical Informatics
- Database Systems
- Computational Intelligence, Methods and Applications
- Discovery Systems in Bioinformatics

Faculty Members
Due to the multidisciplinary nature of the program, the teaching faculty is drawn from the various engineering and science schools of the university, such as the School of Computing Engineering, School of Mechanical and Aerospace Engineering, School of Electrical and Electronic Engineering, School of Chemical and Biomedical Engineering, National Institute of Education, and School of Biological Sciences.

Furthermore, there are several adjunct faculty members from the Genome Institute of Singapore, the National Cancer Centre, and the Institute for Infocomm Research. The advanced BioInformatics Research Centre (BIRC) also provides an interdisciplinary environment and training for students of this program. Currently, these professors are contributing many of the topics in newly emerging areas of bioinformatics. The BIRC is also equipped with a Compaq AlphaServer SC45 supercomputer with 44 nodes (Fig. 1), each node comprising four 1-GHz Alpha processors with 1-GB memory, capable of approximately 300 Gigaflops, supported by 2 TB of storage.

Graduates
Many of the graduates of this program are currently pursuing research in the field; many have managed to do original research from the knowledge gained from this course. Many of them are currently employed in national-level research institutes and related commercial companies. The amount of original research publications in international conferences and journals produced by students in the MSc Bioinformatics program is among the highest of all the MSc programs in Nanyang Technological University, Singapore.