Xin Yao
University of Birmingham, UK

A Research-Led and Industry-Oriented MSc Program in Natural Computation

Natural computation is the study of computational systems that use ideas and get inspirations from natural systems, including biological, physical, chemical, economical and social systems. It covers many active research fields, such as evolutionary computation, neural computation, molecular computation, quantum computation, ecological computation, etc. It has made tremendous progress in recent years in academic research and real-world applications. Many university departments have been offering individual modules/courses in one form or another on related topics. However, few universities have been offering an entire postgraduate program in natural computation. This article summarises one such MSc program in Natural Computation at the University of Birmingham, UK.

The Program

The MSc in Natural Computation at the University of Birmingham, UK, was first launched in the 2001/02 academic year after one year preparation and timely funding from UK's Engineering and Physical Sciences Research Council (EPSRC). The key motivation behind this program is to provide an *integrated* postgraduate program in natural computation to computer science and engineering graduates who will be able to learn different natural computation techniques in a unified framework. The program is designed to be industry-oriented and research-led.

This 12-month Program consists of six taught modules, two mini-projects and one summer project. Among the six taught modules, the three introductory modules, i.e., *Introduction to Evolutionary Computation, Introduction to Neural Computation* and *Introduction to Quantum and Molecular Computation*, are designed to lay a solid foundation for the four major areas within natural computation. In addition to these three taught modules in the first term, the students will do their first mini-project, which consists of a research skill component and a mini-research project. The idea behind the mini-research project, supervised by an academic staff, is to engage students in carrying out literature review, problem formulation, feasibility study and producing preliminary results. The mini-project serves as a precursor to the summer project, for which students are expected to have done the ground work and produce a substantial piece of research work.

The three advanced taught modules, i.e., *Nature Inspired Optimization, Nature Inspired Learning* and *Nature Inspired Design*, are organized very *differently* from other modules. Instead of designing a module around certain techniques, such as evolutionary or neural computation, these three modules are organized around problems, i.e., optimization, learning and design problems. The key message to the students is the following: It is important to learn new technologies, yet it is absolutely essential to understand and learn *how* to best use such technologies and *when* to use them. Guided by such a principle, we have included both conventional and nature-inspired techniques in these modules, e.g., conventional optimization algorithms and statistical machine learning techniques, although the emphasis is on nature-inspired techniques. Students are exposed to the methodology of problem-solving in optimization, learning and design. They are confronted all the time with the question of which algorithm/ technique/ method to use for a given type of problem. There is a strong emphasis on *principled* applications of natural computation techniques, rather than trying one's favorite algorithms on any problem without a good understanding.

Similar to the first term, the students will do their second mini-project in the second term in addition to the three advanced modules. The motivation of having two mini-projects is to encourage (i.e., force) students to broaden their in-depth knowledge in at least two very different areas. The second mini-project must be different from the first one and be supervised by a different academic supervisor. So the students are also exposed to different academic views and styles of supervision. Such diversity and experiences are very beneficial to students in their later career.

The third term of our 12-month program is entirely devoted to the summer research project, which is built upon one of the two mini-projects. Some of these projects are carried out in companies, where the students spend most of their

1556-603X/06/$20.00©2006IEEE

time supervised by an industrial co-supervisor. Separate research reports are required and assessed for all three projects.

Industry Involvement

Industrial involvement is a key feature of this program. From the very beginning, an Industrial Advisory Board was set up with members from BT, Honda, HP, Unilevel, Rolls Royce, e9, etc. The Board meets annually and provides input and feedback to the program in terms of structure and contents.

The industrial partners in the program also provide a lot of other support, e.g., giving guest lectures, proposing and co-supervising student projects, hosting and supporting (financially) students in their companies, sponsoring student prizes, etc. Not only have students benefited from such close industrial involvement, but the industrial partners have also benefited from the research that the students have done and from the access to the pool of excellent students. Some companies have offered jobs to placement students or funded Ph.D. studentships after students' MSc. There have also been joint publications generated from the MSc research projects.

Concluding Remarks

The MSc in Natural Computation at Birmingham, UK, is a 12-month full-time program. It is research-led because of its heavy emphasis on research projects. It is also industry-oriented because of active involvement of industrial partners. Even the basic fundamental research is often inspired by practical applications and grounded on real-world problems. It has a number of studentships available each year to UK and EU students. The program has attracted excellent students from many countries, including UK, USA, China, India, Norway, Greece, Turkey, etc. For more information about the program, please visit **www.cs.bham.ac.uk/study/postgraduate-taught/msc-nc/**. We have a number of fully-funded studentships available each year for UK and EU students.

There are still many challenges in running such a specialist MSc program in universities, which need to be tackled in the future. Firstly, natural computation includes some rather technical and mathematical contents, e.g., in quantum computation and statistical machine learning. Some students have found them difficult. Secondly, natural computation is still known and used mainly by large international companies, which limits students' choices in pursuing their summer project placements within companies. Thirdly, a project-heavy MSc program puts a lot of pressure on staff resource, i.e., the availability of the academic staff in this field and their time. This implies that it would be difficult to have a large number of students in this program.

Acknowledgment

The MSc in Natural Computation has been partially supported by an EPSRC grant (GR/N29969/01) and various industrial partners.

Research Frontier *(continued from page 5)*

Under the assumption that there is one and only one regulatory RNA structure in common to these sequences, "bins" are generated by a random choice of one structure from each list of hits per organism. Each bin is considered as an "individual" in an evolving population of solutions with the goal of maximizing similarity of the bin contents.

Several variation operators including replacement of structure hits (analogous to mutation) and recombination among bins (analogous to crossover) were used. Fitness was an aggregate of components that measure RNA similarity in terms of nucleotide sequence similarity within structural components, structural component length similarity, and similarity of the free energy of the most thermodynamically- stable secondary structures. Tournament selection was used to determine which bins would serve as parents for generation of "offspring bins." Using four dual processor Pentium III 450 mHz computers operating with Linux, search spaces of 1×10^{23} can be searched. An evolving population of 100 parent bins and 50 offspring bins can be used for 21 generations to search 3×10^{-19} of the space in roughly one hour of CPU time and discover solutions that are known to be experimentally verified as true structures that are regulatory features. This same approach can now be used by the pharmaceutical industry to mine RNA sequences for novel regulatory structures that could be useful drug targets.

Conclusions

Computational intelligence approaches have been applied to microarray data to make predictions of diagnosis or prognosis based upon changes in gene expression. Only recently have computational intelligence methods been applied to determine the regulatory features than result in the gene expression observed on the microarray. It is hoped that, by doing so, we may offer insight into the workings of entire systems via the genotype-phenotype mapping and that this understanding will, in turn, result in novel approaches for pharmaceutical discovery.

References

[1] G.B. Fogel, D.G. Weekes, G. Varga, E.R. Dow, H.B. Harlow, J.E. Onyia, and S. Chen, "Discovery of sequence motifs related to coexpression of genes using evolutionary computation," Nucleic Acids Research, vol. 32, no. 13, pp. 3826–3835, 2004.

[2] T.J. Macke, D.J. Ecker, R.R. Gutell, D. Gautheret, D.A. Case, and R. Sampath, "RNAMotif, an RNA secondary structure definition and search algorithm," Nuc. Acids Res., vol. 29, pp. 4724–4735, 2001.

[3] G.B. Fogel, V.W. Porto, D.G. Weekes, D.B. Fogel, R.H. Griffey, J.A. McNeil, E. Lesnik, D.J. Ecker, and R. Sampath, "Discovery of RNA structural elements using evolutionary computation," Nuc. Acids Res., vol. 30, no. 23, pp. 5310–5317, 2002.

[4] E.A. Lesnik, G.B. Fogel, D. Weekes, T.J. Henderson, H.B. Levene, R. Sampath, and D. Ecker, "Identification of Conserved Regulatory RNA Structures in Prokaryotic Metabolic Pathway Genes," BioSystems, vol. 80, no. 2, pp. 145–154, 2004.

[5] G.B. Fogel, D.G. Weekes, R. Sampath, and D.J. Ecker, "Parameter optimization of an algorithm for RNA structure discovery," Proceedings of the 2004 Congress on Evolutionary Computation, IEEE Press, Piscataway, NJ, pp. 607–613, 2004.