Genetic Algorithms in Molecular Design of Novel Fibers

Les Sztandera, leader (PhilaU);
Chih-Chung Chu (Cornell),
Hugh Cartwright (Oxford Univ., UK)

Our goal is twofold: (i) to establish an extensive structure-property correlation database library; (ii) to develop computational techniques that can draw upon data within this library to predict the formulation for a range of novel polymers and fibers with improved characteristics such as exceptional stretch, recovery, strength, increased bulk, improved hand and comfort properties, and which are easy to dye. These technically advanced fibers will combine the best properties of nylon and polyester.

Design Methodology

In the general design process we are using two techniques from within the field of Artificial Intelligence. With a Neural Network, we aim to develop a tool that is able to predict the properties of a given copolymer or homopolymer from its structure, thereby solving what is known as the forward problem. The second technique is the Genetic Algorithm, which solves the inverse problem by forming the heart of the search procedure whose role is to find the optimum formulation.

These two methods are combined together in a collaborative manner, to form a paired algorithm in which the neural network is used in the calculation of the genetic algorithms' fitness function; the tool created in this way is known as a Hybrid Intelligent System. We chose object-oriented design to develop the software tool, rather than a functional approach, because it is considered to produce more maintainable and easily understood system architecture and code.

Initially, our work focused on the second of the two tools required for the Hybrid System, the Genetic Algorithm (GA). The present formulation of the GA comprises two sub-systems: a GA engine with five fitness functions and a graphical user interface (GUI) front end which provides the user with an easy route into the functionality of the GA engine.

The design methodologies we used for the production of the GA are analogous to the evolutionary incremental software engineering model. Starting from a basic genetic algorithm, we added input and output functionalities to yield a checked test-bed model, which we will use as the basis of the polymer modeling system.

We designed a suitable GUI sub-system so the user can enter and modify GA parameters; this incorporates checks to validate parameters as they are entered.

Initial versions of the algorithm included only a single fitness function. In a further development, we added a number of functionalities and incorporated them into the architectures of both the GA engine and the GUI. Four further fitness functions were also incorporated into the GA. Other minor refinements of the GA engine and adjustments in the GUI subsystem have also been made.

We are developing an extensive structure-property database to aid in designing fibers with improved properties, such as stretch, strength, bulk, comfort and dyeability.

Features Dropped at the Design Stage

The Genetic Algorithm is, like many Artificial Intelligence methods, characterized by a substantial degree of flexibility in the way in which it is formulated. To create a productive tool, it is essential that the freedom that exists within the GA algorithm (in both its methodology and selection of parameter values) is fully explored, to find the conditions that are most productive for a given problem. We have therefore been investigating a number of features of the genetic algorithm engine at the design stage.

A key factor in preparing an effective GA model is choosing the nature of the "crossover" step. In uniform crossover, a randomly generated or a specified masking pattern is used in the recombination of the genes of a chromosome. We have investigated various alternative schemes for the crossover. In a typical GA application, the most suitable approach to crossover is strongly correlated with the encoding used to form the GA strings. We therefore anticipate that the method finally adopted for this step will depend upon the method adopted for encoding the structural parameters of the polymer.

We have also considered the best method for the so-called "mating" step in the GA. Typically, this involves the manipulation of two candidates solutions (or "parents"), but it is also possible to use three, four or more parents, or just one. The capability to vary the number of parents has been included as a feature in the GUI to ease this feature's integration in the future.

It is common practice to monitor the average fitness of the genotype population to test for convergence; in other words, to bring the calculation to a halt when the average fitness of a population reaches a certain value or when the fitness fails to improve significantly for a specified number of cycles. There are various ways in which this can be accomplished – at present we assess the mean and median of a sorted set of fitness values; both parameters can indicate when the genotypes in a population converge towards a common solution. The GUI subsystem provides the means to monitor the median genotype fitness.

We are now extending our work to include data from a larger database, and to investigate the prediction of properties of random sets of polymers formed from an alphabet of divalent molecular fragments. This problem has previously been studied using linear correlation; early results suggest that the neural network hybrid method will be a more accurate predictor of polymer properties.

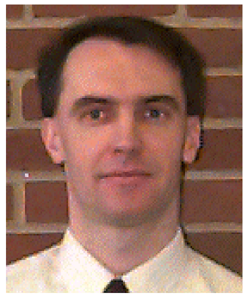
Contributing Graduate Student: *Johnathon Mohr* (PhilaU).

Industry Interactions: []
Other Non-NTC Academic Interactions: ; Government:

Project Web Address:
http://faculty.PhilaU.edu/SztanderaL/NTC/C04-PH02

For Further Information:
1.

Les M. Sztandera, a Professor of Computer Information Systems at PhilaU served as a Distinguished Fulbright FLAD Chair in Information Systems Les earned a Diploma from Cambridge (England) in 1989, an M.S. from Univ. of Missouri in 1990 and a Ph.D. in computer and engineering science from the Univ. of Toledo in 1993. Les' research interests include fuzzy logic, pattern recognition, computer vision, genetic algorithms, neural networks, hybrid intelligent systems, and modeling and management of uncertainty.



I98-P01, S01-PH10, C04-PH02s*, C05-PH1*
sztanderal@philau.edu
(215)-951-5356
http://faculty.philau.edu/sztanderal

Hugh M. Cartwright, a Lecturer and Laboratory Officer in Chemistry at Oxford Univ. (UK), joined the faculty in 1984. He earned a B.Sc. in 1969 and a Ph.D. in 1972 in chemical sciences at the Univ. of East Anglia (Norwich Eng.). Hugh is the author of <u>Applications of Artificial Intelligence in Chemistry</u>. His research interests center on the use of artificial intelligence in scientific problems, such as the dispersal of airborne pollution, optimization of organic synthesis, industrial process control and development, drug design, bacterial growth, bio-informatics and the assessment of medical data.

C04-PH02s, C05-PH01
hugh.cartwright@chem.ox.ac.uk
44 (0) 1865 275 483
http://www.chem.ox.ac.uk/researchguide/hmcartwright.html

Chih Chung Chu, a Professor in Biomedical Engineering at Cornell, joined the faculty in 1978 after 3 years at Univ. of Alabama - Birmingham. C.C. earned a Ph.D. in polymer chemistry from Florida St in 1976 and a B.S. in chemistry Tamkang Univ. (Taiwan) in 1968 and served from 1986-90 as Visiting Research Associate Professor of Surgery at the Hahemann Univ. School of Medicine (Philadelphia). His interests include basic research in polymer/fiber morphology and degradation mechanisms and applied research in biomedical polymers and fibers for human body repair.

M01-CR01*, M03-CR04*, C04-PH02s, C05-PH01
cc62@cornell.edu
(607)-255-1938
http://www.human.cornell.edu/txa/faculty/Chu/whois_frame.html