

**CONNECTING MATHEMATICS AND BIOLOGY IN THE INFORMATION SOCIETY SCHOOLS:
A BRAZILIAN PERSPECTIVE ON TECHNOLOGY USAGE**

Alexandre S. Mendes, FEEC - UNICAMP – Brazil, smendes@densis.fee.unicamp.br

Joni A. Amorim, IMECC - UNICAMP – Brazil, amorimja@yahoo.com

Rosana G. S. Miskulin, FE - UNICAMP – Brazil, rmiskuli@bestway.com.br

Abstract: Technology must be used in mathematics education intending an inclusion of students in the world of contemporary technologies while enhancing mathematics learning. The emphasis in this work is put on the teaching of mathematical and biological sciences with a Brazilian perspective, considering between others the paradigms suggested by the National Curriculum Parameters of the Brazilian Ministry of Education that have as one of their proposals an interdisciplinary approach to teach intending to connect the different fields of study. Artificial intelligence, genetic algorithms, biology, technology in mathematics education and problem solving are the main topics considered.

Introduction

Technology enhances mathematics learning (Amorim and Miskulin, 2001). It is a teaching tool that must be used well by secondary level teachers; these teachers should select and create mathematical tasks that take advantage especially of graphing, visualizing, and computing capabilities through exercises like simulations; this can give students the chance to experience situations that are difficult to create without technological tools.

The emphasis in this work is put on the teaching of Mathematical and Biological Sciences with a Brazilian perspective, considering between others the paradigms suggested by the PCN (National Curriculum Parameters) of the Brazilian Ministry of Education (ME, 1999) that have as one of their proposals an interdisciplinary approach to teach intending to connect the different fields of study. Such considerations are made under the paradigm of a low-cost mass customization (Amorim, 2001) of education in Brazil, fighting the technological illiteracy. This paradigm uses as motivation themes related to the new technologies in a way to help Brazilian students to better adapt to the new Information Society (MCT, 2000) that comes together with the globalization and the Internet.

The above-mentioned PCN intends to give general suggestions to teachers, targeting to enlarge the horizons of students preparing them to a competitive world; the main idea is to show how close are the taught subjects from the everyday lives and communities of these students. This can be achieved through "transversal themes" (ME, 1999): subjects like natural hazards would be considered in the traditional disciplines like mathematics or biology penetrating them during a certain period of classes. The usage of new technologies like genetic algorithms – a branch of the artificial intelligence area – is proposed as a source of transversal themes to motivate students to learn more about science in a way to offer them a chance to become familiar with contemporary technologies.

The applications related to biology and mathematics became part of our lives and can function as an excellent educational thrust through programs at all levels, allowing students not only to understand the world but also preparing them to their future working lives, enhancing their curriculum quality. This work explains, in simple words, the usage of genetic algorithms in problem solving, showing how to connect ideas that are taught in biology and mathematics independently. With a similar approach, teachers from secondary level schools can have one more way to motivate their students with state-of-the-art science examples.

Genetic algorithms, biology, mathematics and problem solving

Complex problems arise in many different areas of knowledge: mathematics, physics, biology, chemistry, economy, etc. Among these, there is a category that is especially difficult: it is called NP-problems. Those problems are object of study of researchers all over the world for the last 30 years, since computers have become widespread.

Complex problem solving

Imagine the following problem: you want to make a trip and want to visit five cities. The tour must start in your city, visit all the five other ones only once and return to the city of origin. In order to know how many possibilities we have, it is necessary to use the concepts of combinatorial analysis. In doing so, we have 120 possibilities (the factorial of 5). Given all the 120 possibilities, which one is the best?

Certainly, the one where the total traveled distance is the smallest, since you don't want to spend all your time and money while on the road, moving between cities. This easy-to-state problem is one of the most difficult in the mathematical literature and brings us to the fundamental concepts of optimization. A problem with five cities can be solved with the help of a road map in less than a minute by hand, just by looking. But with 50 cities, the total number of possible tours raises to $10e^{64}$, a number 65-digit long. But one can say that visiting 50 cities in a single trip is not real. For a vacation trip it really isn't. But consider a work trip, where a single truck must deliver products in several cities. Or worse, suppose a transportation company, with ten trucks transporting a thousand products to one hundred cities. In this scenery, you must divide all the products among the trucks and determine which trucks will visit which cities. After that, you must choose all the ten tours, one for each truck, which minimizes the total traveled distance. Quite complicated, isn't it? With over $10e^{5000}$ possible solutions, this is also a NP-problem, very common in medium and large companies.

The NP-problems share one characteristic: they have too many solutions, a few good, some not so good and many really bad. There is also one, or just a few, called optimal, with no other solution being better than they are. They are the ones we are searching for and as they are only a few in a universe of "zillions" of solutions, finding them is like finding a needle in a haystack. There are several ways of searching such optimal solutions. Some are plain stupid, like making the computer try all the possibilities, picking the best. For most problems, such trial-and-error strategy would take simply too long to finish. Other approaches are much clever and among them we can find the genetic algorithms.

Biology and evolution

Darwin was an English naturalist and when he joined the crew of the HMS Beagle he was only 22 years old. The travel took five years and the ship went to places like Brazil, Argentina and Chile. But, his research was marked by the features he saw in the Galápagos Islands. Islands are special places from the biological point of view. Due to their isolation, species evolve with little external interference. The Galápagos is an archipelago with dozens of small islands and a rich natural life. Darwin found in some of these islands species of birds, with their beak adapted to the most common type of fruit in each island. The species were very similar, except for their beaks, which varied a lot. He concluded that probably all birds belonged to the same species when they arrived in the islands, but adaptation to the many types of food led to a differentiation among them after several hundred years of evolution.

The main mechanism of evolution is natural selection. According to it, the fitter individuals have more chances to survive and generate descendents, due to competition – competition for food, to escape from predators, etc. During the adaptation process, birds with many types of beaks might have appeared, but only the birds with the beak best suited for the fruit present in each island have survived. The others could not find adequate food and got extinct. That was a strong indicative of a natural selection mechanism.

"It is not the strongest of the species that survive, nor the most intelligent, but the one most responsive to change....I can see no limit to the amount of change, to the beauty and complexity of the co-adaptations between all organic beings, one with another and with their physical conditions of life, which may have been affected in the long course of time through nature's power of selection, that is by the survival of the fittest" (Darwin, 1859).

What are the evolution mechanisms that created so many different species? Differentiation occurs through two main mechanisms: genetic recombination and mutation. All complex living beings, from ants to humans, have what is called a genetic code composed by a set of chromosomes; molecules with very complex chemical structures. In humans, they define everything in a person: height, color of eyes, skin, type of hair, predisposition for certain diseases, general corporal conformation, etc. All these characteristics are chemically coded into 23 pairs of chromosomes.

Genetic recombination occurs by mixing the information present in the chromosomes of the parents. The child's genetic code, basically, is created blending parts of the chromosomes of the father and the mother. This is the main reason why genetically favored parents tend to generate also genetically favored children. In general species, individuals with good features tend to live longer, generating more descendents, which will inherit those features.

“Lastly, isolation, by checking immigration and consequently competition, will give time for any new variety to be slowly improved; and this may sometimes be of importance in the production of new species” (Darwin, 1859).

The second mechanism of evolution is mutation. It is a random change in a part, or parts, of the genetic code. As all genetic-level processes are chemically controlled, such changes are induced, usually by mistake, by accidental chemical reactions. The random nature makes the mutation generate features that can be good or bad, but since the genetic code of any complex being is too intricate, random changes tend to destroy information. In fact, over 90% of mutations are destructive, creating worse or non-viable individuals. Such individuals are usually eliminated through natural selection and the results of disastrous mutations are not perpetuated. But the good side is that, when successful, mutation can create good features that are so complex that it could take thousands of years of recombination and natural selection to obtain the same result “naturally”, in the absence of mutation.

Genetic Algorithms

Researchers try to simulate nature on computers since the 1960's, when computers started to get into the universities and research centers. The genetic algorithm is nothing more than a way to simulate the evolution of a species in a computer. It belongs to a branch of research called Artificial Intelligence (AI), which is a computer science discipline that utilizes natural-phenomena-related methods to solve complex problems. GAs are evolution simulators that try to find good solutions for complex problems through evolution itself. The first impulse of the GAs occurred in 1975, after a researcher called John Holland published the book “Adaptation in Natural and Artificial Systems”. Holland started his studies in the area in the beginning of the 1960's but his book marked the boom.

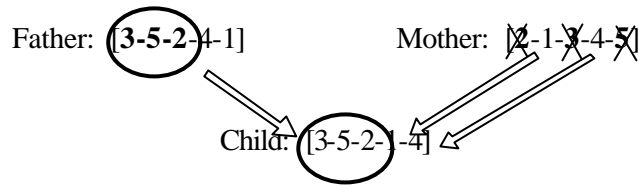
After this brief introduction let's consider how to make the bridge between mathematics and biology, through genetic algorithms. First, what is an algorithm? An algorithm is a method. A sequential, structured method to do something. For example, a cake recipe is an algorithm to make food, or a cake in this case. So, we conclude that genetic algorithms are methods, based on biological concepts, to solve complex problems. Genetic algorithms work in a very similar way evolution works with species. But instead of evolving species, genetic algorithms evolve solutions for the problem. The objective is to evolve individuals (solutions) of a species considering the specific environment (the problem to be solved) so in the end of the evolutionary process, the remaining individuals will be highly adapted (high quality, or even optimal solutions for the specific problem). How do we make this link individuals/solutions and environment/problem?

First, we must define a chromosome, or a genetic code for the solutions of the problem. Take as example that problem of finding the optimal tour when one wants to visit a set of cities, explained earlier. What would be a good representation for a solution? If the cities are numbered, from 1 to n , where n is the number of cities, a simple permutation representation arises, being necessary just one chromosome. Take an example considering 5 cities: an individual with a chromosome configuration of [3-5-2-4-1] means the first city to be visited will be the number three, followed by cities number five, two, four and finally city number one. This genetic code can represent all the possible solutions for the mathematical problem and, of course, the optimal one. That defines how the solutions for our complex problem can be genetically coded into individuals of a species.

The next step is to define a recombination strategy, or how two individuals can create a child. A very interesting recombination is the Order Crossover, proposed by a researcher called Goldberg, in 1989. Let's see how it works using an example. Initially take two parents representing two solutions (individuals) of the problem – a father and a mother.

Father: [3-5-2-4-1] Mother: [2-1-3-4-5]

The Order Crossover starts selecting at random a piece of the chromosome from the father. Suppose that the piece [3-5-2] was selected. This piece is copied into the child, that becomes [3-5-2-*-*], where * are cities not yet determined. These cities will be determined using information from the tour represented by the mother. The mother's chromosome begins with city 2, which is already present in the child. Since the tour can not have repeated cities, go to the next. The next is city number one and as it is not yet present in the child, it should be copied. The child now becomes [3-5-2-1-*]. The next city in the mother is 3, already present in the child. Then comes city 4, which is also copied, completing the child, which becomes [3-5-2-1-4]. Next we show a diagram of the example.



This recombination scheme has a very important feature found in biological species recombination. The child always inherits characteristics present in both parents, being similar to them in many aspects. Now, a mutation scheme is necessary to finish defining the two evolutionary mechanisms that act at the chromosome level. As said earlier, mutation generates a small change in the individual's genetic code. For our test problem, a convenient mutation is the swap of cities. It works in a simple way: two cities are chosen at random and they swap their positions in the chromosome. See the example next.



In the example, cities 4 and 5 were chosen to be mutated and swap their positions. In a problem with only five cities, changing two of them represents an inadequately hard mutation (40% of the genetic code was changed), but for larger problems it becomes more realistic.

With the two genetic-level mechanisms well defined, natural selection is the next thing to be defined. Natural selection depends on the individual's adaptation to the environment. As the environment is the problem itself – to find the smallest possible tour – an adaptation measure is the tour length. Individuals representing small tours can be considered more adapted than individuals representing larger tours. The last thing to define is how to make the species evolve, step by step. A simple genetic algorithm should follow the steps:

1. **Create a population of random tours:** They will be your starting points. As these individuals are usually randomly created, their adaptation will be very low. But do not worry: they will improve. A reasonable size for the population can be around 100 individuals. In nature, species are composed by thousands or millions of individuals, but when we simulate nature in computers, it is not possible to use such large populations due to technical limitations.
2. **Select individuals for recombination:** In most species, all individuals can generate descendents. For the genetic algorithm, it seems fair to select, say, 50% of the population for recombination. If the population is composed of 100 individuals, select 25 pairs of parents. This selection can be made at random. Originally, Holland proposed a selection that benefited the best individuals, selecting them more frequently. But for educational purposes, a random selection, without any bias, seems perfectly adequate.
3. **Create the new individuals:** Each pair of parents will generate one child using the recombination strategy explained earlier. Twenty-five new individuals will be created at this step.
4. **Mutate the new individuals:** Select 5 individuals from the 25 created in **step 3** and mutate them using the mutation strategy described before. This selection can be made at random. Why choosing just 5 out of 25? In nature, mutation is generally very light, or else important information acquired during millions of years of evolution can be lost. Five individuals out of 25 seem to be very reasonable, but the reader can try other values and see how the population evolves.
5. **Substitute the less adapted individuals by the new ones:** Select the 25 less adapted individuals of the population and substitute them by the 25 new ones created in **steps 3 and 4**. In this part of the algorithm the natural selection takes place. In nature, the worse individuals usually die faster, being constantly replaced by the recently born ones.
6. **Return to step 2:** This step closes what is so-called a generation. A generation is a cycle of evolution: new individuals are born, a part of the population dies and gets replaced. As the generations keep evolving, a noticeable improvement of the length of the tours present in the population should take place.

The genetic algorithm described before is very simple, but it works, indeed. In the last 25 years, researchers have proposed several refinements to Holland's initial algorithm, making it much more powerful. Nonetheless, we still cannot guarantee that the optimal solution of the problem being solved is to be always found. But it is much better, and realistic, to have a very good solution in a few seconds, or minutes, than wait years for the best one.

Conclusion

"It is not just the role that mathematics, science, and technology play in the changing economy and workplace that matters. Mathematics and science have become so pervasive in daily life that we tend to overlook them. Literacy in these areas affects the ability to understand weather and stock reports, develop a personal financial plan, or understand a doctor's advice. Taking advantage of mathematical and scientific information does not generally require an expert's grasp of those disciplines. But it does require a distinctive approach to analyzing information. We all have to be able to make accurate observations, develop conjectures, and test hypotheses — in short, we have to be familiar with a scientific approach" (NCMST, 2000).

As previously said, researchers try to simulate nature on computers since the 1960's, when computers started to get into the universities and research centers. In the world present societies, computers are finally turning to be part of the daily lives of students at the secondary level. This work shows how to make the bridge between mathematics and biology through genetic algorithms; this interdisciplinary approach to teach intending to connect the different fields of study can be a way to prepare students to be able to make accurate observations, develop conjectures, and test hypotheses. Contemporary technologies are teaching tools that must be used well by secondary level teachers. In this sense, students can be motivated with advanced concepts once they are treated in simple ways. Further work could consider the preparation of a site with a tutorial on genetic algorithms in the way suggested in this paper. This tutorial could be prepared collaboratively and should be accessible to any one interested through Internet browsers (Amorim and Miskulin, 2001).

Acknowledgments

This work was supported by "Fundação de Amparo à Pesquisa do Estado de São Paulo" (FAPESP – Brazil).

References

- Amorim, J. A. (2001) 'Customizando a Educação Matemática na Realidade Brasileira: Custos e Possibilidades', *III Simposio de Educacion Matematica*, <http://www.edumat.com.ar>, Argentina.
- Amorim, J. A. and Miskulin, R. G. S. (2001) 'A Sociedade da Informação: Uma Possível Cooperação em Busca de Soluções Ibero-Americanas para Educação Matemática', *IV Congresso Ibero-Americano de Educação Matemática*, Bolivia.
- Darwin, C. (1859). 'On the Origin of Species'.
- Ministério da Ciência e Tecnologia (MCT) (2000). *Sociedade da Informação no Brasil - Livro Verde*, Brazil.
- Ministério da Educação (ME) (1999). *Parâmetros Curriculares Nacionais - Ensino Médio*, Brazil.
- The National Commission on Mathematics and Science Teaching for the 21st Century (NCMST) (2000). *Before It's Too Late - A Report to the Nation from The National Commission on Mathematics and Science Teaching for the 21st Century*, <http://www.ed.gov/americaaccounts/glenn/report.doc>, U.S.A.
- United Nations Educational, Scientific and Cultural Organization (UNESCO/IFIP) (1994). *Informatics for Secondary Education - A Curriculum for Schools*, <http://www.unesco.org>.