Zbigniew Michalewicz, *University of Adelaide, AUSTRALIA*, and Matthew Michalewicz, *SolveIT Software, AUSTRALIA*

Machine Intelligence, Adaptive Business Intelligence, and Natural Intelligence

Machine Intelligence

Most people recognize Larry Fogel for his work on evolutionary programming; today, evolutionary programming is considered one of the early branches of evolutionary algorithms, together with genetic algorithms, evolution strategies, genetic programming, and many other—sometimes unnamed—population-based techniques (Baeck et al., 1997). However, it is important to remember that Larry's main interest at that time was in machine intelligence, and his work on evolutionary programming was just to address some issues of machine intelligence.

One of the key observations of Larry Fogel was that machine intelligence might be defined as the capability of a system to adapt its behavior to meet desired goals in a range of environments. Consequently, intelligent behavior requires prediction. An additional component of intelligent behavior is adaptation (which is based on prediction, as adaptation to future circumstances requires predicting those circumstances). The final component of intelligent behavior is the capability of taking appropriate action (Fogel et al., 1966). Consequently, the foundation for evolutionary programming research was to generate machine intelligence by simulating the evolutionary process on a class of predictive algorithms (as opposed to the approach of generating machine intelligence by replicating humans by directly creating rules to follow or creating the neural connections).

It is interesting to observe that Larry Fogel identified three key elements of intelligence, namely:

❏ ability to predict,
❏ ability to adapt, and
❏ ability to take appropriate action.

Clearly, there is no need to argue that prediction is important—without this capability no system (including natural systems) can be called intelligent. The concept of adaptability is certainly gaining popularity. Adaptability has already been introduced in everything from automatic car transmissions (which adapt their gear-change patterns to a driver's driving style), to running shoes (which adapt their cushioning to a runner's size and stride), to Internet search engines (which adapt their search results to a user's preferences and prior search history). These products are very appealing to individual consumers, because, despite their mass production, they are capable of adapting to the preferences of each unique user after some period of time. Finally, the ability to take appropriate action is probably the most important component of an intelligent system. After all, if a system can predict and adapt, but all the decisions made

© DIGITALSTOCK

are silly, the system hardly deserves to be called "intelligent!" Note also, that the term "appropriate action" implies optimization, as usually the system should take (or recommend) the *best* course of action.

Interestingly, the three components of prediction, adaptation, and optimization constitute the core modules of adaptive business intelligence systems. When we discuss adaptive business intelligence in the next section of the paper, the connection with evolutionary programming will become apparent.

One additional aspect of Larry Fogel's research was connected with the concept of so-called "Valuated State Space"®. The "Valuated State Space"® approach provides a convenient structure for assessing various decision-making parameters in terms that people are familiar with. Further, it allows individuals to apply subjective relative importance weights and provides a mechanism for dealing with degrees of criticality of parameters. The "Valuated State Space"® approach provides a rank ordering of all possible outcomes and rapid comparison of two or more potential decisions to determine which is better.

For a high-level overview of valuated state spaces the reader is referred to (Michalewicz and Fogel, 2004); however,

it is worthwhile to look at this approach as a general problem-solving approach. After all, the major steps in "Valuated State Space"® imply some general problem-solving steps, which require understanding of the problem, rejection of intuition, building a model of the problem by defining the variables, constraints, and the objectives. Thus the systematic approach proposed for the Valuated State Spaces can be translated into a general problem solving methodology. When we discuss Puzzle-Based Learning in the third section of the paper, the connection with "Valuated State Space"® will be visible.

These links with Larry Fogel's work define the organization of this paper. The next part of the paper presents the main concepts behind adaptive business intelligence, and the following part discusses the current state of a new approach to learning, called Puzzle-Based Learning. A short section on Larry's and authors' business experience concludes the paper.

Adaptive Business Intelligence

Since the computer age dawned on mankind, one of the most important areas in information technology has been that of "decision support." Today, this area is more important than ever. Working in dynamic and ever-changing environments, modern-day managers are responsible for an assortment of far reaching decisions: *Should the company increase or decrease its workforce? Enter new markets? Develop new products? Invest in research and development?* The list goes on. But despite the inherent complexity of these issues and the ever-increasing load of information that business managers must deal with, all these decisions boil down to two fundamental questions:

❏ What is likely to happen in the future?
❏ What is the best decision right now?

Whether we realize it or not, these two questions pervade our everyday lives—both on a personal and professional level. When driving to work, for instance, we have to make a traffic prediction before we can choose the quickest driving route. At work, we need to predict the demand for our product before we can decide how much to produce. And before investing in a foreign market, we need to predict future exchange rates and economic variables. It seems that regardless of the decision being made or its complexity, we first need to make a prediction of what is likely to happen in the future, and then make the best decision based on that prediction. This fundamental process underpins the basic premise of adaptive business intelligence.

Simply put, adaptive business intelligence is the discipline of combining prediction, optimization, and adaptability into a system capable of answering these two fundamental questions: *What is likely to happen in the future? And, what is the best decision right now?* (Michalewicz et al. 2007). To build such a system, we first need to understand the methods and techniques that enable prediction, optimization, and adaptability (Dhar and Stein, 1997). At first blush, this subject matter is nothing new, as hundreds of books and articles have already been written on business intelligence (Vitt et al, 2002; Loshin, 2003), data mining and prediction methods (Weiss and Indurkhya, 1998; Witten and Frank, 2005), forecasting methods (Makridakis et al., 1988), optimization techniques (Deb 2001; Coello et al. 2002; Michalewicz and Fogel, 2004), and so forth. However, none of these has explained how to combine these various technologies into a software system that is capable of predicting, optimizing, and adapting. Adaptive business intelligence addresses this very issue.

Clearly, the future of the business intelligence industry lies in systems that can make decisions, rather than tools that produce detailed reports (Loshin 2003). As most business managers now realize, there is a world of difference between having good knowledge and detailed reports, and making smart decisions. Michael Kahn, a technology reporter for Reuters in San Francisco, makes a valid point in his January 16, 2006 story titled, "Business Intelligence Software Looks to Future":

"But analysts say applications that actually answer questions rather than just present mounds of data is the key driver of a market set to grow 10 percent in 2006 or about twice the rate of the business software industry in general.

'Increasingly you are seeing applications being developed that will result in some sort of action,' said Brendan Barnacle, an analyst at Pacific Crest Equities. 'It is a relatively small part now, but it is clearly where the future is. That is the next stage of business intelligence.'"

Business Intelligence vs. Adaptive Business Intelligence

"The answer to my problem is hidden in my data … but I cannot dig it up!" This
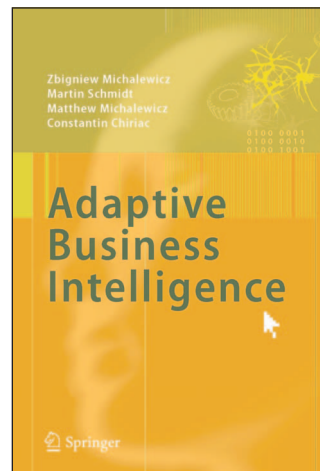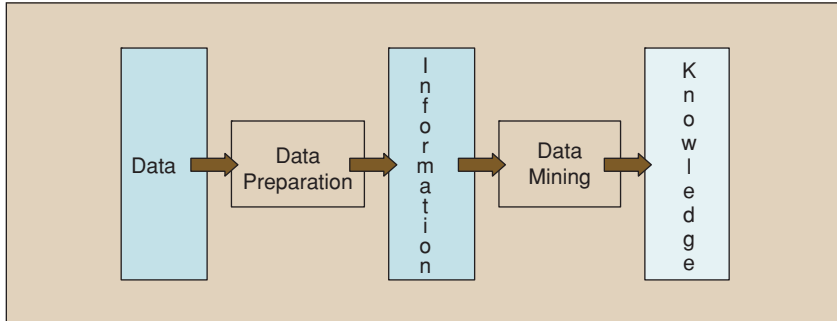
FIGURE 1 The process of Business Intelligence.

popular statement has been around for years as business managers gathered and stored massive amounts of data in the belief that they contain some valuable insight. But business managers eventually discovered that raw data are rarely of any benefit, and that their real value depends on an organization's ability to analyze them. Hence, the need emerged for software systems capable of retrieving, summarizing, and interpreting data for end-users (Moss and Atre, 2003).

This need fueled the emergence of hundreds of business intelligence companies that specialized in providing software systems and services for extracting knowledge from raw data. These software systems would analyze a company's operational data and provide knowledge in the form of tables, graphs, pies, charts, and other statistics. For example, a business intelligence report may state that 57 percent of customers are between the ages of 40 and 50, or that product X sells much better in Florida than in Georgia.[1]

Consequently, the general goal of most business intelligence systems was to: (1) access data from a variety of sources; (2) transform these data into information, and then into knowledge; and (3) provide an easy-to-use graphical interface to display this knowledge. In other words, a business intelligence system was responsible for collecting and digesting data, and presenting knowledge in a friendly way (thus enhancing the end-user's ability to make good decisions). The following diagram illustrates the processes that underpin a traditional business intelligence system:

Although different texts have illustrated the relationship between data and knowledge in different ways (e.g., Davenport and Prusak, 2006; Prusak, 1997;

[1]Note that business intelligence can be defined both as a "state" (a report that contains knowledge) and a "process" (software responsible for converting data into knowledge).
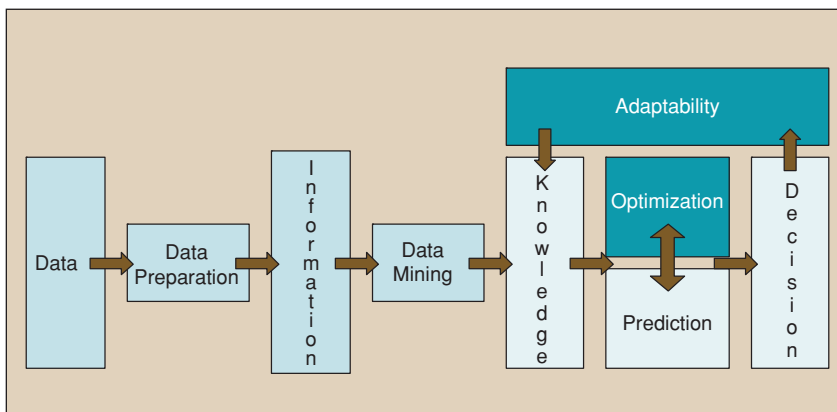
Shortliffe and Cimino, 2006), the commonly accepted distinction between data, information, and knowledge is:
- ❏ *Data* are collected on a daily basis in the form of bits, numbers, symbols, and "objects."
- ❏ *Information* is "organized data," which are preprocessed, cleaned, arranged into structures, and stripped of redundancy.
- ❏ *Knowledge* is "integrated information," which includes facts and relationships that have been perceived, discovered, or learned.

Because knowledge is such an essential component of any decision-making process (as the old saying goes, *"Knowledge is power!")*, many businesses have viewed knowledge as the final objective. But it seems that knowledge is no longer enough. A business may "know" a lot about its customers—it may have hundreds of charts and graphs that organize its customers by age, preferences, geographical location, and sales history - but management may still be unsure of what decision to make! And here lies the difference between "decision support" and "decision making:" all the knowledge in the world will not guarantee the right or best decision.

Moreover, recent research in psychology indicates that widely held beliefs can actually hamper the decision-making process. For example, common beliefs like "the more knowledge we have, the better our decisions will be," or "we can distinguish between useful and irrelevant knowledge," are not supported by empirical evidence. Having more knowledge merely increases our confidence, but it does not improve the accuracy of our decisions. Similarly, people supplied with "good" and "bad" knowledge often have trouble distinguishing between the two, proving that irrelevant knowledge decreases our decision-making effectiveness.

Today, most business managers realize that a gap exists between having the right knowledge and making the right decision. Because this gap affects management's ability to answer fundamental business questions (such as

FIGURE 2 The process of Adaptive Business Intelligence.

"What should be done to increase profits? Reduce costs? Or increase market share?"), the future of business intelligence lies in systems that can provide answers and recommendations, rather than mounds of knowledge in the form of reports. *The future of business intelligence lies in systems that can make decisions!* As a result, there is a trend emerging in the marketplace called adaptive business intelligence. In addition to performing the role of traditional business intelligence (transforming data into knowledge), Adaptive business intelligence also includes the decision-making process, which is based on prediction and optimization:

While business intelligence is often defined as "a broad category of application programs and technologies for gathering, storing, analyzing, and providing access to data," adaptive business intelligence can be defined as "the discipline of using prediction and optimization techniques to build self-learning 'decisioning' systems" (as the above diagram shows). Adaptive business intelligence systems include elements of data mining, predictive modeling, forecasting, optimization, and adaptability, and are used by business managers to make better decisions.

This relatively new approach to business intelligence is capable of recommending the best course of action (based on past data), but it does so in a very special way: An adaptive business intelligence system incorporates prediction and optimization modules to recommend near-optimal decisions, and an "adaptability module" for improving future recommendations. Such systems can help business managers make decisions that increase efficiency, productivity, and competitiveness. Furthermore, the importance of adaptability cannot be overemphasized. After all, what is the point of using a software system that produces sub par schedules, inaccurate demand forecasts, and inferior logistic plans, time after time? Would it not be wonderful to use a software system that could *adapt* to changes in the marketplace? A software system that could *improve* with time?

Current Trends

Adaptability is a vital component of any intelligent system, as it is hard to argue that a system is intelligent if it does not have the capacity to adapt. For humans, the importance of adaptability is obvious: our ability to adapt was a key element in the evolutionary process. In the case of artificial intelligence, consider a chess program capable of beating the world chess master: Should we call this program intelligent? Probably not. We can attribute the program's performance to its ability to evaluate the current board situation against a multitude of possible "future boards" before selecting the best move. However, because the program cannot learn or adapt to new rules, the program will lose its effectiveness if the rules of the game are changed or modified. Consequently, because the program is incapable of learning or adapting to new rules, the program is not intelligent.

The growing popularity of adaptability is also underscored by a recent publication of the U.S. Department of Defense. This lists 19 important research topics for the next decade and many of them include the term "adaptive:" *Adaptive* Coordinated Control in the Multi-agent 3D Dynamic Battlefield, Control for *Adaptive* and Cooperative Systems, *Adaptive* System Interoperability, *Adaptive* Materials for Energy-Absorbing Structures, and Complex *Adaptive* Networks for Cooperative Control.

For sure, adaptability was recognized as an important component of intelligence quite some time ago: Alfred Binet (born 1857), French psychologist and inventor of the first usable intelligence test, defined intelligence as "...judgment, otherwise called good sense, practical sense, initiative, the faculty of *adapting* one's self to circumstances." Adaptability is a vital component of any intelligent system, as it is hard to argue that a system is "intelligent" if it does not have the capacity to adapt. For humans, the importance of adaptability is obvious: our ability to adapt was a key element in the evolutionary process. In psychology, a behavior or trait is adap-tive when it helps an individual adjust and function well within a changing social environment.

The same holds true for any expert system. No one questions the usefulness of expert systems in some environments (which are usually well defined and static), but expert systems that are incapable of learning and adapting should not be called "intelligent." Some expert knowledge was programmed in, that is all.

So, what are the future trends for adaptive business intelligence? In words of Jim Goodnight, the CEO of SAS Institute (Collins et al. 2007):

"Until recently, business intelligence was limited to basic query and reporting, and it never really provided that much intelligence..."

However, this is about to change. Keith Collins, chief technology officer of SAS Institute (Collins et al. 2007) believes that:

"A new platform definition is emerging for business intelligence, where BI is no longer defined as simple query and reporting. ... In the next five years, we'll also see a shift in performance management to what we're calling predictive performance management, where analytics play a huge role in moving us beyond just simple metrics to more powerful measures."

Further, Jim Davis, vice president of marketing at SAS Institute (Collins et al. 2007) stated:

"In the next three to five years, we'll reach a tipping point where more organizations will be using BI to focus on how to optimize processes and influence the bottom line..."

Research Issues

Every problem has an objective. Usually, this is a general statement describing what we are looking for. The objective defines the goal (or set of goals) for a particular problem. These goals are translated into evaluation functions, which provide mappings from the solution space to a set of numbers.

Thus, evaluation functions assign numeric values for each solution for each specified goal.

Evaluation functions (for single-objective problems) or a set of evaluation functions (for multi-objective problems) are key components of any heuristic method (whether genetic algorithms, tabu search, simulated annealing, ant system, or even simple hill-climbers), as they define the connection between the method and the problem. By assigning a numeric quality measure to each solution, evaluation functions allow comparison between the qualities of various candidate solutions. Note that evaluation functions may return just the rank of a candidate solution among a set of solutions, a precise number (when the evaluation function is defined as a closed formula), or they may include various components (as penalty expressions for cases when a candidate solution violates some problem-specific constraints).

Many real world problems are set in uncertain (possibly changing) environments. There is a general agreement (Jin & Branke, 2005) that such uncertainties can be categorized into four classes: (1) noise, (2) robustness, (3) approximation, and (4) time-varying environments. Consequently, evaluation functions should be modified accordingly to deal with each particular case. However, it seems that the above classification misses the most important (and probably most frequent) real world scenario: namely, where the evaluation functions are based on predictions of the future values of some variables. Before we present and discuss the fifth category, and argue that this fifth category is the most common in real word situations, let's discuss the main features of these four categories.

Noise

Sometimes evaluation functions are subject to noise. This happens when evaluation functions return sensory measurements or results of randomized simulations. In other words, the evaluation procedure for the same solution (i.e., the solution defined as a vector of some design variables) may return different values. The common approach in such scenarios is to approximate a noisy evaluation function *eval* by an averaged sum of several evaluations:

$$eval(\mathbf{x}) = 1/n \Sigma_{i=1\ldots n} f(\mathbf{x}) + z_i,$$

where $\mathbf{x}$ is a vector of design variables (i.e., variables controlled by a method), $f(\mathbf{x})$ is the evaluation function, z_i represents additive noise, and n is the sample size. Note that the only measurable (returned) values are $f(\mathbf{x}) + z$.

Robustness

Sometimes design variables, other variables, or constraints of the problem are subject to perturbations after the solution is determined. The general idea is that such (slightly modified) solutions should have quality evaluations (thus making the original solution robust). This is important in scenarios involving manufacturing tolerances, or when it is necessary to modify the original solution because of employee illness or machine failure. The common approach to such scenarios is to use evaluation function eval based on the probability distribution of possible disturbances δ, which is approximated by Monte Carlo integration:

$$eval(\mathbf{x}) = 1/n \Sigma_{i=1\ldots n} f(\mathbf{x} + \delta_i).$$

Note that $eval(\mathbf{x})$ depends on the shape of $f(\mathbf{x})$ at point $\mathbf{x}$; in other words, the neighborhood of $\mathbf{x}$ determines the value of $eval(\mathbf{x})$.

Approximation

Sometimes it is too expensive to evaluate a candidate solution. In such scenarios, evaluation functions are often approximated based on experimental or simulation data (the approximated evaluation function is often called the meta-model). In such cases, evaluation function *eval* becomes:

$$eval(\mathbf{x}) = f(\mathbf{x}) + E(\mathbf{x}),$$

where $E(\mathbf{x})$ is the approximation error of the meta-model. Note that the approximation error is quite different than noise, as it is usually deterministic and systematic.

Time-Varying Environments

Sometimes evaluation functions depend on an additional variable: time. In such cases, evaluation function *eval* becomes:

$$eval(\mathbf{x}) = f(\mathbf{x}, t),$$

where t represents the time variable. Clearly, the landscape defined by the function f changes over time; consequently, the best solution may change its location over time. There are two main approaches for handling such scenarios: (1) to restart the method after a change, or (2) require that the method is capable of chasing the changing optimum.

However, it seems the largest class of real world problems is not included in the above four categories. From our business/industry experience of the last decade, it is clear that in many real world problems the evaluation functions are based on the predicted future values of some variables. In other words, evaluation function *eval* is expressed as:

$$eval(\mathbf{x}) = f(\mathbf{x}, P(\mathbf{x}, \mathbf{y}, t)),$$

where $P(\mathbf{x}, \mathbf{y}, t)$ represents an outcome of some prediction for solution vector $\mathbf{x}$ and additional (environmental, beyond our control) variables $\mathbf{y}$ at time t. Let's compare this category with the four categories defined earlier to see the differences between them.

First of all, noise may or may not be involved. If the prediction model is deterministic, then there is no noise in the scenario: every solution vector x is evaluated to the same value. On the other hand, if the prediction model involves simulations, noise might be present. Second, the meaning of robustness is quite different. Unexpected disturbances (e.g., delays) influence the outcomes of the prediction model, and should be handled accordingly. Third, the concept of approximation is different. Note, that in some cases we can evaluate a candidate solution

precisely (e.g., when the evaluation function is not expensive), however, approximation is connected with uncertainties of the predictions. Finally, the time-changing environment also has a different meaning. As the real world changes, the prediction model needs constant updates and/or parameter adjustments, thus changing the problem landscape in an implicit way.

There are some recent, successful implementations of adaptive business intelligence systems reported (e.g., Michalewicz et al. 2005), which provide daily decision support for large corporations and result in multi-million dollars return on investment. There are also companies (e.g., www.solveitsoftware.com) that specialize in development of adaptive business intelligence tools. However, further research effort is required. For example, most of the research in machine learning has focused on using historical data to build prediction models. Once the model is built and evaluated, the goal is accomplished. However, because new data arrive at regular intervals, building and evaluating a model is just the first step in adaptive business intelligence. Because these models need to be updated regularly (something that the adaptability module is responsible for), we expect to see more emphasis on this updating process in machine learning research. Also, the frequency of updating the prediction module, which can vary from seconds (e.g., in real-time currency trading systems), to weeks and months (e.g., in fraud detection systems), may require different techniques and methodologies. In general, adaptive business intelligence systems would include the research results from control theory, statistics, operations research, machine learning, and modern heuristic methods, to name a few. We also expect that major advances will continue to be made in modern optimization techniques. In the years to come, more and more research papers will be published on constrained and multi-objective optimization problems, and on optimization problems set in dynamic environments. This is essential, as most real world business problems are constrained, multi-objective, and set in a time-changing environment.

Natural Intelligence

What is missing in most curricula—from elementary school all the way through to university education—is coursework focused on the development of problem-solving skills. Most students never learn how to think about solving problems—throughout their education, they are constrained to concentrate on specific questions at the back of textbooks. So, without much thinking, they apply the material from each chapter to solve a few problems given at the end of each chapter (why else would a problem be at the end of the chapter?). With this type of approach to "problem solving," it is unsurprising that students are ill prepared for framing and addressing real world problems. When they finally enter the real world, they suddenly find that problems do not come with instructional textbooks.

Although many educators are interested in teaching "thinking skills" rather than "teaching information and content," the fact remains that young people often have serious difficulties in independent thinking (or problem-solving skills) regardless of the nature of a problem. As Alex Fisher wrote in his book, "Critical Thinking:"

"… though many teachers would claim to teach their students 'how to think,' most would say that they do this indirectly or implicitly in the course of teaching the content, which belongs to their special subject. Increasingly, educators have come to doubt the effectiveness of teaching 'thinking skills' in this way, because most students simply do not pick up the thinking skills in question."

This approach has dominated the educational arena—whether in history, physics, geography, or any other subject—almost ensuring that students never learn how to think about solving problems in general.

The Approach

Over the past few decades, various people and organizations have attempted to address this educational gap by teaching "thinking skills" based on some structure (e.g. critical thinking, constructive thinking, creative thinking, parallel thinking, vertical thinking, lateral thinking, confrontational and adversarial thinking). However, all these approaches are characterized by a departure from mathematics as they concentrate more on "talking about problems" rather than "solving problems." It is our view that the lack of problem solving skills in general is the consequence of decreasing levels of mathematical sophistication in modern societies.

Hence, we believe that a different approach is needed. To address this gap in the educational curriculum, we have created a new course (based on our new book, "Puzzle Based Learning: An Introduction to Critical Thinking, Mathematics, and Problem Solving") that focuses on getting students to *think* about how to frame and solve *unstructured* problems (those that are not encountered at the end of some textbook chapter …). The idea is to increase the student's mathematical awareness and problem solving skills by discussing a variety of *puzzles*. In other words, we believe that the course should be based on the best traditions introduced by Gyorgy Polya and Martin Gardner during the last 60 years. In one of our favorite books, "Entertaining Mathematical Puzzles," Martin Gardner wrote:

"Perhaps in playing with these puzzles you will discover that mathematics is more delightful than you expected. Perhaps this will make you want to study the subject in earnest, or less hesitant about taking up the study of a science for which a knowledge of advanced mathematics will eventually be required."

Many other mathematicians have expressed similar views. For example, Peter Winkler in his book "Mathematical Puzzles: A Connoisseur's Collection," wrote: "*I have a feeling that understanding and appreciating puzzles,*

> **Interestingly, the three components of prediction, adaptation, and optimization constitute the core modules of adaptive business intelligence systems.**

even those with one-of-a-kind solutions, is good for you."

This new course is the result of many years of experience in educating young engineers, mathematicians, and computer scientists on many levels at many universities in many countries (USA, Mexico, Argentina, New Zealand, Australia, South Korea, Japan, China, Poland, Sweden, Germany, Spain, Italy, France and the UK). Limited experiments using the puzzle-based learning approach with these students have already produced outstanding course evaluations and countless comments that praise the problem-solving orientation of the course. We believe that the main reasons behind most students' enthusiasm for the puzzle-based learning approach are:

❏ Puzzles are educational, as they illustrate many useful (and powerful) problem-solving rules in a very entertaining way.
❏ Puzzles are engaging and thought-provoking.
❏ Contrary to many textbook problems, puzzles are not attached to any chapter (as is the case with real world problems).
❏ It is possible to talk about different techniques (e.g. simulation, optimization), disciplines (e.g. probability, statistics), or application areas (e.g. scheduling, finance) and illustrate their significance by discussing a few simple puzzles. At the same time, the students are aware that many conclusions are applicable to the broader context of solving real world problems.

Some Supporting Evidence

As a matter of fact, the puzzle-based learning approach has a much longer tradition than just 60 years. The first mathematical puzzles were encountered in Sumerian texts that date back to around 2,500 BC! Yet the best evidence of the puzzle-based learning approach can be found in the works of Alcuin, an English scholar born around A.D. 732 whose main work was "Problems to Sharpen the Young"—a text which included over 50 puzzles. Some twelve hundred years later, one of his puzzles is still used by countless artificial intelligence textbooks![2]

The first author is a member of the editorial board of the "International Journal Teaching Mathematics and Computer Science." It is clear that new methods of teaching (especially engineers) are sought and experimented with. Further, one of the earlier books by the first author, "How to Solve It: Modern Heuristics" (written together with Larry's son, David), included a selection of puzzles to illustrate some problem solving activities. Despite the fact that the book aimed at graduate students interested in genetic algorithms, neural networks, fuzzy systems, and many other traditional and modern techniques, the readers—because of puzzles—got much more than just information of particular techniques. Some comments (still available on www.amazon.com) were:

❏ "This book teaches you how to think of a solution for the problem you face…"
❏ "…anyone interested in […] human thinking should read and understand this book."
❏ "I used this book in a Master's class on heuristics (Systems Engineering,

University of Virginia) and received the most positive textbook reviews I have seen in my fifteen years of teaching."
❏ "Most importantly, it does so in a way that no other book I've seen does—it makes it fun and it makes you *think*!"

Importance of Mathematics

Over the years, two primary approaches to problem solving have emerged. One is the *technical* approach (represented in many textbooks), which concentrates on specific problem-solving techniques. The other is the *psychological* approach, which is based on structural thinking—meaning that some structure is imposed on the thinking process during the problem solving activity.

Let's discuss these two approaches in a bit more detail; for that purpose we have selected two popular texts. The first one is "Operations Research: An Introduction," by Hamdy A. Taha, and the other is a book by Edward de Bono, "Six Thinking Hats." The first book illustrates the technical approach very well, as it is loaded with mathematical techniques for a variety of different problems. On the other hand, the second book presents a particular structured way of thinking. Let us have a closer look at these two books.

"Operations Research: An Introduction," by Hamdy A. Taha consists of several chapters, each of which relate to a specific problem type. For example, there is a chapter on linear programming, which is a particular technique for solving problems with many variables and where the objective and the values of these variables are expressed as linear expressions. Another chapter of Taha's book discusses a transportation model and its variants, while another presents a series of techniques applicable to network models. There are chapters on goal programming, integer linear programming, dynamic programming, inventory models, forecasting models, etc. Each chapter includes selected references and a problem set. For example, the chapter on inventory models includes the following exercise:

---

[2]The puzzle is the "river crossing problem" (we will return to this puzzle in chapter 12 of this book): A man has to take a wolf, a goat, and some cabbage across a river. His rowboat has enough room for the man plus either the wolf or the goat or the cabbage. If he takes the cabbage with him, the wolf will eat the goat. If he takes the wolf, the goat will eat the cabbage. Only when the man is present are the goat and the cabbage safe from their enemies. All the same, the man carries wolf, goat, and cabbage across the river. How has he done it?

"McBurger orders ground meat at the start of each week to cover the week's demand of 300 lb. The fixed cost per order is $20. It costs about $0.03 per lb per day to refrigerate and store the meet. (a) Determine the inventory cost per week of the present ordering policy. (b) Determine the optimal inventory policy that McBurger should use, assuming zero lead time between the placement and receipt of an order. (c) Determine the difference in the cost per week between McBurger's current and optimal ordering policy."

Clearly, the problem is well-defined and very specific. Earlier parts of the chapter on inventory models discussed, of course, a general inventory model (where the total inventory cost is given as a total of purchasing cost, setup cost, holding cost, and shortage cost) and the classic economic order quantity models. The formula is derived in the chapter to provide the optimum value of the order quantity y (number of units) as a function of setup cost K associated with the placement of an order (in dollars per order), demand rate D (in units per time unit), and holding cost h (in dollars per inventory unit per time unit). The model suggests to order:

$$y = \sqrt{2KD/h}$$

units every y/D time units. Again, it is not our goal to scare you by providing a formula in the introductory part of this text (especially that the derivation of this formula requires some calculus …) but rather to point out the specific nature of the problem and the specific (and very precise) solution. This is a good illustration of the technical approach.

It seems that Taha's text is similar to many other texts from disciplines such as engineering, mathematics, finance, and business, in that it has two main characteristics:

(a) The problem types and corresponding techniques are very specific; and
(b) mathematics is used extensively.

However, there is usually no discussion on "how to solve a problem"—the text gives some recipes on how to arrive at a solution once the problem has already been reduced to the problem type defined in the text. As indicated in the preface, students are constrained to concentrate on textbook questions at the back of each chapter, using the information learned in that chapter … So all these specialized texts (whether on probability, statistics, simulations, etc.) that represent the technical approach for problem solving, do not present a problem-solving methodology. They just provide (very useful) information of particular techniques for particular classes of problems.

Let us now turn our attention to the other book, Edward de Bono's "Six Thinking Hats," which represents the psychological approach. As we have indicated earlier, the book suggests some structure for the thinking process during the problem solving activity. In particular, each of six hats represents a particular function of the thinking process:

White Hat—collection of objective facts and figures
Red Hat—presentation of emotional view
Black Hat—discussion of weaknesses in an idea
Yellow Hat—discussion on benefits of the idea
Green Hat—generation of new ideas
Blue Hat—imposition of control of the whole process

The general idea is that instead of thinking simultaneously along many directions, a thinker should do one thing at the time. Edward de Bono explains it very clearly:

"The main difficulty of thinking is confusion. We try to do too much at once. Emotions, information, logic, hope and creativity all crowd in on us. It is like juggling with too many balls.

What I am putting forward in this book is a very simple concept, which allows a thinker to do one thing at a time. He or she becomes able to separate emotion from logic, creativity from information, and so on. The concept is that of the six thinking hats. Putting on any one of these hats defines a certain type of thinking."

It seems that "Six Thinking Hats" is characterized by two facts (as are many other texts on thinking processes, which include texts on critical thinking, constructive thinking, creative thinking, parallel thinking, vertical thinking, lateral thinking, confrontational, and adversarial thinking, to name a few):

(a) the problem types and corresponding "techniques" are not very specific. The approach is very general and it applies to most problems (as opposed to specific problem types); and
(b) the approach is mathematics-free.

Indeed, the examples given in the "Six Thinking Hats" vary from house selling activities, through advertising and marketing issues, to pricing products. Further, the mathematics is nonexistent despite the fact that some problems may require more precise mathematics. There is no question that the approach proposed by Edward de Bono is very useful and that many corporations benefited from the methodology of "Six Thinking Hats." On the other hand, the rejection of mathematics in "Six Thinking Hats" expresses itself even in the author's statements, such as: *"In a simple experiment with three hundred senior public servants, the introduction of the Six Hats method increased thinking productivity by 493 percent."* Well, this is very impressive, but any person with any "critical thinking" skills (or

> **What is missing in most curricula—from elementary school all the way through university education—is coursework focused on the development of problem-solving skills.**

> **Clearly, the future of the business intelligence industry lies in systems that can make decisions, rather than tools that produce detailed reports.**

some fancy for precision) may ask for clarifications:

> What is the definition of productivity (especially in cases of senior public servants) and how such productivity—and improvement in productivity—has been measured?

Indeed, these are very important questions. In the case of the public servants, did three hundred employees fill out forms that evaluated their (increased) productivity? If so, then this can be compared to an example provided by Darrell Huff in his book "How to Lie with Statistics." The San Francisco Chronicle published an article titled "British He's Bathe More Than She's," and the story supported the title with the following facts (based on hot-water survey of 6,000 representative British homes). "The British male over five years of age soaks himself in a hot tub on an average of 1.7 times a week in the winter and 2.1 times in the summer. British women average 1.5 baths a week in the winter and 2.0 in the summer." Darrell Huff, discussing this case, made an excellent (and very important) observation. He wrote:

> "…the major weakness is that the subject has been changed. What the Ministry really found out is how often these people said they bathed, not how often they did so. When a subject is an intimate as this one is, with the British bath-taking tradition involved, saying and doing may not be the same after all."

It seems that the same argument can be applied to the public servants … Most likely, their productivity was measured in hours (i.e., the shorter the time to make a decision, the better). Edward de Bono explains:

> "A major corporation used to spend twenty days on their multi-national project team discussion. Using the parallel thinking of the Six Hats method, the discussions can now take as little as two days."

However, if this was that case, then it seems there is something fundamentally very wrong with the whole picture, as *the quality* of the decisions reached is completely ignored and not measured! We should not care so much whether the problem solving process took x or y hours, as *the quality* of solution is the most important aspect.

There is an excellent book (on science and education, one can say) by Eliyahu M. Goldratt and Jeff Cox, "The Goal." The book describes the struggle of a plant manager who tries to improve factory performance. He worries about productivity, excess inventories, throughput, balancing capacities, and many other measurements. Only with the help of a consultant does he realize that there is only one goal and one measurement: *"The goal of a manufacturing organization is to make money and everything else we do is means to achieve the goal."* Similarly, in the problem solving process there is only one goal: To find the best possible solution. Of course, very often there is a trade-off between the time needed to find a solution and the quality of the solution (this is often discussed in computer science courses on analysis of algorithms), but is seems that the "Six Thinking Hats" method is concerned with only the secondary aspect of problem solving: time efficiency. Precise evaluation of the solution is of lesser importance.

Thus the psychological approach looks like the *opposite extreme* of the technical approach in the spectrum of problem-solving methodologies, as the former focuses on organizational issues of "thinking" for general problems, rather than specific techniques on how to arrive at a solution. Furthermore, the psychological approach uses natural language to describe its mechanisms, whereas the technical approach uses mathematics as a problem solving language.

Which of these two approaches (technical versus psychological) should be used in the real world? Well, each of these two approaches has a crowd of enthusiasts and supporters; however, it seems that the technical approach is based on the solid fundamentals of science. Even some philosophers and psychologists tend to agree. One of the pearls of wisdom taught by Anthony de Mello in his famous book, "One Minute Wisdom," was the following observation:

> "Weeks later, when a visitor asked him what he taught his disciples, he said, 'To get their priorities right: Better have the money than calculate it; better have the experience than define it.'"

It is easy to extend the above statements (while preserving their spirit) by stating that:

> It is better to know how to solve problems than to have the ability to talk about them!

On the other hand, representatives of the technical approach admit that "*although mathematics is a cornerstone of Operations Research, one should not 'jump' into using mathematical models until simpler approaches have been explored. In some cases, one may encounter a 'commonsense' solution through simple observations. Indeed, since the human element invariably affects most decision problems, a study of the psychology of people may be key to solving the problem*," (Hamdy A. Taha, "Operations Research: An Introduction"). These comments are followed by a delightful example, where the problem of slow elevator service in a large office building was solved not by the use of mathematical queuing analysis or simulation, but by installing full-length mirrors at the entrance to the elevators: the complaints disappeared as people were kept occupied watching themselves (and others) while waiting for the elevator!

Clearly, there are many merits in concepts related to critical, vertical, lateral, and other thinking paradigms. However, mathematics—the queen of all sciences—must remain the universal language of problem solvers. Otherwise, as we saw, there is a danger of making imprecise statements, and what is worse, there is a danger of finding—and implementing—poor solutions!

Numerous mathematicians have put a lot of effort into finding a middle ground between the technical and psychological approaches to problem solving. The best known work, without a doubt, is Gyorgy Polya's "How to Solve It," which stands out as one of the most important contributions to problem-solving literature of the twentieth century. Even now, as we have moved into the new millennium, the book continues to be a favorite among teachers and students for its instructive methods. Other works include "I Hate Mathematics," written by Marylin Burns, which is full of tips and methods for solving problems.

Current State

Our new course (which aims at getting engineering students to *think* about how to frame and solve *unstructured* problems) has been approved by the University of Adelaide for faculty of engineering, computer science, and mathematics (altogether seven schools). The course will be offered in two versions: (a) full-semester course and (b) a unit within a general course (e.g., Introduction to Engineering). Many other universities are in the preliminary phase of introducing such a course. All teaching materials (power point slides, assignments) are being prepared. The new

textbook ("Puzzle Based Learning: An Introduction to Critical Thinking, Mathematics, and Problem Solving") will be available July 2008.

We believe that besides being a lot of fun, the puzzle-based learning approach will also do a remarkable job of convincing engineering students that (a) science is useful and interesting, (b) the basic courses they take are relevant, (c) mathematics is not that scary (no need to hate it!), and (d) it is worthwhile to stay in school, get a degree, and move into the real world, which is loaded with interesting problems (problems perceived as real world puzzles…). These points are important, as most students are unclear about the significance of the topics covered during their studies. Oftentimes, they do not see a connection between the topics taught (e.g. linear algebra) and real world problems, and they lose interest with predictable outcomes.

Conclusions

This paper has presented some thoughts on how Larry Fogel has impacted the authors' lives in a variety of different ways. However, this paper would not be complete without an additional observation. Larry Fogel was one of few scientists who also created a business to implement his ideas in the real world. Indeed, the authors of this paper have done the same, by starting and selling out of a company in the United States, and a few years later establishing another company in Australia. The authors have described their business experiences in their recent book: "Winning Credibility: A guide for building a business from rags to riches," (see *www.WinningCredibility.com*).
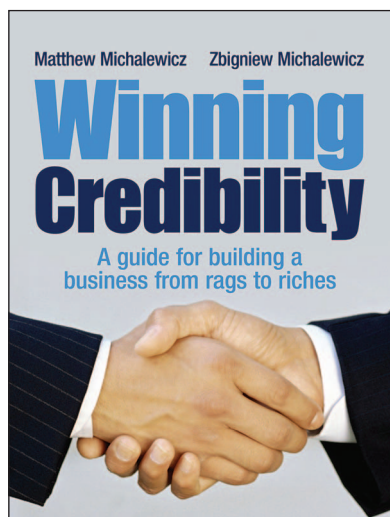
References

[1] T. Baeck, D. Fogel, and Z. Michalewicz, (Editors), *Handbook of Evolutionary Computation*, joint publication of Oxford University Press and Institute of Physics, New York and London, 1997.

[2] M. Burns, *I Hate Mathematics*, Little, Brown, and Company, New York, 1975.

[3] C.A.C. Coello, A.A. Van Veldhuizen, and G.B. Lamont, *Evolutionary Algorithms for Solving Multi-Objective Problems*, Kluwer Academic, 2002.

[4] K. Collins, J. Goodnight, M. Hagström, J. Davis, J., *The Future of Business Intelligence: Four Questions, Four Views*, SASCOM, First quarter, 2007.

[5] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*, Wiley, 2001.

[6] E. de Bono, *Six Thinking Hats*, MICA Management Resources, New York, 1999.

[7] A. de Mello, *One Minute Wisdom*, Doubleday, New York, 1986.

[8] V. Dhar and R. Stein, *Seven Methods for Transforming Corporate Data into Business Intelligence*, Prentice Hall, 1997.

[9] A. Fisher, *Critical Thinking: An Introduction*, Cambridge University Press, Cambridge, 2001.

[10] L.J. Fogel, A.J. Owens, and M.J. Walsh, *Artificial Intelligence through Simulated Evolution*, Wiley, New York, 1966.

[11] Martin Gardner, *Entertaining Mathematical Puzzles*, Dover Publications, New York, 1961.

[12] M. Gardner, *My Best Mathematical and Logic Puzzles*, Dover Publications, New York, 1994.

[13] D. Huff, *How to Lie with Statistics*, W. W. Norton, New York, 1993.

[14] Y. Jin J. and Branke, "Evolutionary optimization in uncertain environments—A survey," *IEEE Transactions on Evolutionary Computation*, vol. 9, no. 3, pp. 303–317, June 2005.

[15] D. Loshin, *Business Intelligence: The Savvy Manager's Guide*, Morgan Kaufmann, 2003.

[16] S. Makridakis, S.C. Wheelwright, and R.J. Hyndman, *Forecasting: Methods and Applications*, Wiley, 1998.

[17] Z. Michalewicz and D.B. Fogel, *How to Solve It: Modern Heuristics*, 2nd edition, Springer, Berlin, 2004.

[18] M. Michalewicz and Z. Michalewicz, *Winning Credibility: A guide for building a business from rags to riches*, 2nd edition, Hybrid Publishers, Melbourne, Australia, 2007.

[19] Z. Michalewicz and M. Michalewicz, *Puzzle Based Learning: An Introduction to Critical Thinking, Mathematics, and Problem Solving*, Hybrid Publishers, Melbourne, 2008.

[20] Z. Michalewicz, M. Schmidt, M. Michalewicz, and C. Chiriac, "A Decision-support system based on computational intelligence: A case study," *IEEE Intelligent Systems*, vol. 20, no. 4, pp. 44–49, 2005.

[21] Z. Michalewicz, M. Schmidt, M. Michalewicz, and C. Chiriac, *Adaptive Business Intelligence*, Springer, Berlin, 2007.

[22] L.T. Moss and S. Atre, *Business Intelligence Roadmap*, Addison Wesley, 2003.

[23] J.A. Paulos, *Innumeracy: Mathematical Illiteracy and Its Consequences*, Hill and Wang, New York, 1988.

[24] G. Polya, *How to Solve It: A New Aspect of Mathematical Method*, Princeton University Press, Princeton, 1945.

[25] H.A. Taha, *Operations Research: An Introduction*, 7th edition, Pearson Education, Upper Saddle River, 2003.

[26] S.M. Weiss and N. Indurkhya, *Predictive Data Mining*, Morgan Kaufmann, 1998.

[27] E. Vitt, M. Luckevich, and S. Misner, *Business Intelligence: Making Better Decisions Faster*, Microsoft Press, 2002.

[28] I.H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd edition, Morgan Kaufmann, 2005.

[29] P. Winkler, *Mathematical Puzzles: A Connoisseur's Collection*, A.K. Peters, Wellesley, 2004.