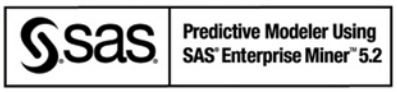
Business Intelligence
Solutions

Head Office
T: 267.616.1444
F: 1.905.761.1006
9 Flagstaff Place
Philadelphia, PA
19115 USA

info@bisolutions.us

Toronto
T: 647.271.1932
F: 905.761.1006
150 Borrows Street
Thornhill, ON
L4J 2W8 Canada

www.bisolutions.us



www.BISolutions.us | Data. Knowledge. Action.

WHITE PAPER:

REGRESSION ANALYSIS AND WICKED BUSINESS PROBLEMS

JUNE 2008

Eugene Brusilovskiy & Dmitry Brusilovsky

The Concept of the Wicked Business Problem

The complexity of the fundamental business problems has been increasing dramatically over last several decades. In most cases, the traditional steps ('understand the problems', 'gather information,' 'synthesize information,' 'work out solutions') are no longer a good option due to the fact that problems often have incomplete, contradictory, and changing requirements, and due to complex interdependencies, it is difficult to find the appropriate solutions. While attempting to solve such a problem (the so called wicked problem), the solution of one of its aspects may reveal or create other, even more complex problems.

According to Rittel and Webber (1), Horn and Weber (2), and Conklin (3) a wicked problem has the following features:

1. There is no definitive formulation of a wicked problem, and it is not understood until after formulation of a solution.
2. Wicked problems have no stopping rule.
3. Solutions to wicked problems are not true-or-false, but better or worse.
4. There is no immediate and no ultimate test of a solution to a wicked problem.
5. Every solution to a wicked problem is a "one-shot operation"; because there is no opportunity to learn by trial-and-error, every attempt counts significantly.
6. Each attempt to create a solution changes the understanding of the problem.
7. Wicked problems do not have an enumerable (or an exhaustively describable) set of potential solutions, nor is there a well-described set of permissible operations that may be incorporated into the plan.
8. Every wicked problem is essentially unique, but there is no unique "correct" view of the problem.
9. Every wicked problem can be considered a symptom of another problem.
10. The existence of a discrepancy representing a wicked problem can be explained in numerous ways. The choice of explanation determines the nature of the problem's resolution.
11. Stakeholders have radically different world views and different frames for understanding the problem.
12. The problem has considerable uncertainty, ambiguity, and it is never solved.
13. The problem often has political, economic, and cultural constraints

Therefore, finding a solution of a wicked problem is more art than science.

It is not clear from the beginning what the problem is and thus, what a solution might be. The processes of solving and specifying the wicked problem develop in parallel and drive each other. Since the problem cannot be defined, it is difficult to tell when it is resolved. There are no unambiguous criteria for deciding if the problem is resolved, and getting all stakeholders to agree that a resolution is 'good enough' can be a challenge. In other words, the solutions found are often such that they still could be improved and it is up to the problem solver to decide when enough is enough.

Objectives

Often, a business problem can be represented as a three-component expression $Y = F(X)$, where X is the input information, Y is the output, and F is a mechanism of transforming the input into the output. In a well-defined business problem it is possible given, two of these components, find the absent one:

1. Induction: given Y and X, find F
2. Deduction: given X and F, find Y
3. Abduction: given F and Y, find X

With a well-defined business problem, it is clear what the problem is and the solution can be clearly specified as well. The criteria for assessing the quality of a solution can be determined when the solution is obtained.

The Well-Defined Regression Problem as a Model for Well-Defined Data-Driven Induction Problem

A well defined business induction problem can be converted into a well-defined data-driven regression problem, $Y = F(X)$, where X is the given vector of independent variables (inputs), Y is the given dependent variable (output) and F is an unknown mechanism (that has to be estimated) of transforming the given inputs X into the given output Y.

The mechanism F can be learned from the data (e.g., non-parametric regression: TreeNet, CART, Neural Net, Lowess (local regression), Generalized Additive Model (GAM), Multivariate Adaptive Regression Splines (MARS), Random Forest, etc.). In simple cases, the mechanism F can be set up by a researcher (linear regression, non-linear parametric regression). There are plenty of fit statistics to evaluate the accuracy of a model on unseen data when the model is developed.

What is a Wicked (Ill-Defined) Business Problem?

A wicked (or ill-defined) business problem has two unknown components in the expression *Y = F(X)*. In other words, given one component we need to find the other two. There are three possibilities:

1. Given *X*, find *F* and *Y*
   Risk-based pricing in the banking industry (4)is a good example of a type 1 ill-defined business problem. The bank will offer a better (lower) interest rate to the borrower whose likelihood of default is thought to be lower. A financial analyst knows the borrower attributes X (credit scores from different bureaus, average salary, employment duration, etc.). The problem is then to estimate the lending risk associated with the consumer (it is a collateral goal) along with the corresponding price Y of a loan (final goal). The relationship F between known consumer attributes X and unknown loan price Y is also unknown.

2. Given *F*, find *X* and *Y*
   The Customer Life Time Value (CLTV) calculation can be treated as an example of this type of ill-defined business problems. Unknown customer attributes such as average customer lifetime, average cost of customer retention, etc., along with product, company, and market attributes represent the unknown *X* information, while the CLTV represents the unknown *Y* information. What is known in this situation is the mechanism *F* of transforming the customer/product/company/ market attributes into the CLTV.

3. Given *Y*, find *X* and *F*
   The Workers' Compensation Insurance Fraud Detection problem (5) can serve as an example of a type 3 ill-defined business problem. A fraud investigator knows *Y* – that is, the claim cost, claim duration, disability duration, and other "inflated" claim characteristics. The problem is to estimate the unknown real claim characteristics *X* and how they turned into the "inflated" claim characteristics *Y*, i.e. to identify the fraud mechanism *F*.

The Wicked Regression Problem as a Model for a Wicked Data-Driven Problem. An Iterative Approach to the Solution.

From now on, we will assume that all three components *X, Y,* and *F* can be only approximately identified, using the available data, subject matter knowledge, expert judgment and the intuition of the decision maker. Then an approximate relationship can be written in the following way: $y \sim f(x)$, where *x, y* and *f* are rough guesses of *X, Y,* and *F* respectively, and the symbol '~' reflects the fact that both sides of the expression are only in rough agreement. Rough guesses of *x, y,* and *f* are the first step in the iterative process of finding a reasonable/acceptable solution. The goal of the next iterations is to combine all three types of decisions – induction, deduction and abduction to "improve" the new estimates of the unknown *X, Y* and *F*. The problem here is that there are no objective criteria to judge the quality of new estimates. As we mentioned earlier, only the synergy of the decision maker's intuition, data-induced knowledge and expert judgment can produce meaningful updates. Non-parametric and non-linear regression analysis can help estimate the relationship *F*. The incorporation of expert judgment into the modeling process is described in (6).

Conclusion

Failure to recognize that a problem at hand is a wicked one essentially complicates the process of understanding the problem, the definition of a solution, and the search for an acceptable solution.
If the problem can be formulated as a wicked data-driven regression problem, then an iterative process of combining inductive, deductive and abductive paradigms can be used to improve the convergence of the finding and its correctness. Evaluating the quality of a model and settling on a solution is more art than science.

References

1. Rittel, Horst, and Melvin Webber (1973), *Dilemmas in a General Theory of Planning,* Policy Sciences, 4:155-169.
2. Horn, Robert E., and Robert P. Weber (2007), *New Tools For Resolving Wicked Problems: Mess Mapping and Resolution Mapping Processes*, Strategy Kinetics L.L.C.
3. Conklin, Jeff (2006), *Wicked Problems & Social Complexity*, Chapter 1 of Dialogue Mapping: Building Shared Understanding of Wicked Problems, Wiley.
4. Brian King, *Risk-Based Pricing*. Available at www.benchmarkinternational.com
5. Terry Woodfield (2006), *Predicting Workers' Compensation Insurance Fraud Using SAS Enterprise Miner 5.1 and SAS Text Miner*. Available at http://www2.sas.com/proceedings/sugi30/071-30.pdf
6. P.M. Brusilovskiy, L.M. Tilman (1996), *Incorporating Expert Judgment Into Multivariate Polynomial Modeling*, Decision Support Systems, 18:199-214.

