Interdisciplinary Journal of Information, Knowledge, and Management Volume 4, 2009

Discovering a Decision Maker's Mental Model with Instance-Based Cognitive Mining: A Theoretical Justification and Implementation

David M. Steiger and Natalie M. Steiger
University of Maine, Orono, Maine, USA

dsteiger@maine.edu; nsteiger@maine.edu

Abstract

The purpose of this paper is to provide a theoretical justification for, and describe an implementation of, instance-based cognitive mining (ICM), a process that analyzes multiple decision instances using the inductive learning algorithms of artificial intelligence to generate a mathematical representation of the decision maker's mental models, *explicitly* relating how the decision maker *implicitly* selects and weighs key factors in making decisions within a specific problem domain. The foundation and justifications of ICM are based on three distinct literatures: 1) knowledge creation (mental models and knowledge externalization), 2) cognitive science (tacit knowledge and instance-based learning), and 3) artificial intelligence (data mining and inductive learning networks). We also propose an architecture that integrates several technologies to capture and express a decision maker's mental model, and we develop a prototype ICM software implementation.

Finally, we describe a preliminary experiment that applies the ICM process to small teams of decision makers that tests (and supports) two hypotheses: H1: the ICM-derived mental model representation provides the mediating causal process through which the set of key factor values affects the actual decisions made by a team of decision makers; and H2a: the ICM-derived mental model representation is consistent with the team's self-reported algebraic, directional, or tacit relationship(s), and the team's self-reported key factors; or H2b: any significant differences between the ICM-derived mental model representation and the team's self-reported relationships or key factors are not consistent with the team's actual decisions.

Keywords: cognitive science, cognitive mapping, decision support systems, knowledge management, externalization, inductive learning algorithms, mental models, tacit knowledge.

Introduction

A decision maker's knowledge involves both explicit and tacit knowledge (Choo, 1998; Davenport & Prusak, 1998; Sun, Merrill & Peterson., 2001). Explicit knowledge is defined as knowledge that can be expressed formally and can, therefore, be easily communicated or diffused throughout an organization (Choo, 1998). In decision making, explicit knowledge can be found in trade journals, executive speeches, written procedures, etc. On the other hand, tacit knowledge consists of subjective exper-

Editor: Zlatko Kovacic

tise, assumptions, and insights that an individual develops from being immersed in an activity or profession for an extended period of time (Choo, 1998; Reber, 1989). Tacit knowledge is integrated and stored in the form of mental models that become so ingrained in the decision maker's thought processes that they are instinctive and thus not easily verbalized or communicated (Kearney & Kaplan, 1997; Tsoukas, 2003). Once formed, mental models provide the decision maker with a method of filtering and processing data and information within a specific decision domain, generating potential alternatives, simulating the outcomes of those alternatives, and evaluating the outcomes with respect to an appropriate set of criteria (Johnson-Laird, 1983; Senge, 1990; Weick, 1990); i.e., mental models form the basis of decision making. More generally, mental models provide the internal representation of reality to the individual (Nonaka & Konna, 1998) and become the center of all knowledge creation, learning, and (potentially) improved decision making. As such, they are the critical component of knowledge creation, a theory which requires that tacit mental models be made known explicitly; however, there exists minimal theoretical or practical direction for implementing this requirement.

The purpose of this paper, which is also its contribution to the literature, is to provide a theoretical justification for, and describe an implementation of, instance-based cognitive mining (ICM), a process that analyzes multiple decisions made by an individual (or small team of individuals) in a specific decision domain, using inductive learning algorithms and regression to generate a mathematical representation of the decision maker's mental model, *explicitly* relating how the decision maker *implicitly* selects and weighs key factors in making decisions. The foundation and justifications for ICM are based on three distinct literatures: 1) knowledge creation (mental models and knowledge externalization), 2) cognitive science (tacit knowledge and instance-based learning), and 3) artificial intelligence (data mining and inductive learning networks). We also propose an architecture that integrates several technologies to discover and measure mental models and we develop a prototype ICM implementation. Note that the ICM process is applicable to both individual decision makers and small teams of decision makers; in the latter, the team must form a consensus for each decision situation. This consensus forming requirement encourages team members to share both tacit and explicit knowledge during the decision process, a complementary "externalization" process as suggested in Nonaka's (Nonaka & Takeuchi, 1995) knowledge spiral.

Finally, we describe a preliminary experiment that applies the ICM process to small teams of decision makers that tests (and supports) two hypotheses: H1 – the ICM-derived mental model representation provides the mediating causal process through which the set of key factor values affects the actual decisions made by a team, with no *a priori* knowledge of which factors were actually used in its decisions, algebraic relations between these key factors (e.g., cross terms, squared terms, inverse relationships, etc.), or relative weighting of the factors; and H2a – the ICM-derived mental model representation is consistent with the team's self-reported algebraic, directional, or tacit relationships, and the self-reported key factors; or H2b – any significant differences between the ICM-derived mental model representation and the team's self-reported relationships or key factors are not consistent with the team's actual decisions.

ICM can be classified as a technique within the extended theory of cognitive mapping. In this theory, cognitive mapping is defined as the process of capturing and describing the important features of an individual's mental model in a specific decision domain, including the assumptions, key factors (conceptual, logical, physical), and their interrelationships (temporal, causal, spatial, mathematical) (Kearney & Kaplan, 1997). Current cognitive mapping techniques span a wide variety of assumptions, inputs, processing, and outputs, and have been categorized, reviewed, and compared by several researchers (e.g., Carley & Palmquist, 1992; Hodgkinson, 1997; Huff, 1990; Jonassen & Grabowski, 1993; Mohammed, Klimoski & Rentsch, 2000; Walsh, 1995). ICM is unique when compared to existing cognitive mapping techniques in each of three basic characteristics; specifically, ICM's input is a set of decision instances, its primary processing technique includes inductive learning networks, and its output is a mathematical (perhaps nonlinear) representation of the relationships between the decisions and the associated situations represented by key factor values.

The paper is organized as follows. In the next section, we provide the theoretical foundation and justifications for ICM as a process for discovering a decision maker's mental model. In the third section, we describe the ICM process, in general, and our software implementation, in particular. In the fourth section, we describe a preliminary test of ICM and provide the test results. In the fifth section, we present a summary of our research and propose several areas for future research.

Theoretical Justification of Instance-based Cognitive Mining

The theoretical justifications for ICM's foundation and assertions are based on three distinct literatures: knowledge creation, cognitive science, and artificial intelligence. Each of these literatures also suggests important characteristics that ICM's implementation should incorporate.

Theoretical Justification from Knowledge Creation

Nonaka's theory of knowledge creation (Nonaka & Takeuchi, 1995) views tacit knowledge and explicit knowledge as complementary entities and suggests that there is a knowledge spiral, beginning and ending with the decision maker, that can create, amplify, and crystallize new knowledge. This knowledge spiral consists of sharing tacit knowledge between individuals (socialization), converting tacit knowledge into explicit knowledge (externalization), integrating this explicit knowledge with other explicit knowledge (combination), and forming, updating, and/or enhancing the mental model held in the mind of the decision maker (internalization). The knowledge spiral is driven by the conversion of tacit knowledge (i.e., the decision maker's mental model) into explicit knowledge so that it can be shared with and analyzed by others and combined with other explicit knowledge to improve decision making. To ignore this tacit-to-explicit conversion process is to cripple the individual's and firm's knowledge creation process and stifle organizational learning. That is, externalizing a decision maker's mental model is "the quintessential knowledge creation process" (Choo, 1998, p. 122), the very heart of improved decision making (Davenport & Prusak, 1998, p. 8), and the key to enhancing the long-term competitive advantage of the firm (Nonaka & Takeuchi, 1995). And yet, mental model externalization is the most problematic stage of the knowledge spiral. Except for a general listing of platitudes (Nonaka & Konna, 1998, p. 44) that have been found ineffective to date (Becerra-Fernandez & Sabherwal, 2001, p. 48), this all important step is largely unaddressed in terms of implementing Nonaka's theory.

For example, Becerra-Fernandez and Sabherwal (2001), in an empirical study focusing on perceived knowledge satisfaction in a technical organization (NASA's Kennedy Space Center), found that although knowledge externalization significantly influenced perceived knowledge satisfaction, externalization experienced very low levels of usage. Their conclusion was that "greater efforts are necessary to build tools and techniques that facilitate (knowledge) externalization" (Becerra-Fernandez & Sabherwal, 2001, p. 48). Further, "externalization processes require new approaches, unlike internalization processes that benefit from teaching and training methods" (Sabherwal & Becerra-Fernandez, 2003, p. 249). These conclusions form a major justification for our research and development efforts in ICM.

Theoretical Justification from Cognitive Science

In cognitive science literature, instance-based learning has been a staple of research for the last 25+ years (Gonzales, Lerch, & Lebiere., 2003; Simon & Gobet, 1996). An instance is defined as a triplet consisting of a set of task-relevant factors, a solution or decision based on those factors, and the outcome of the decision and its utility to the decision maker (Gonzales et al., 2003); different situations are represented by different values of the factors that describe the situation, leading to (potentially) different decisions. Decision makers learn to focus on those instance factors that they judge to be most relevant, ignoring other irrelevant factors (Simon & Gobet, 1996), combining those key factor values and the decision maker's mental model to make decisions (Johnson-Laird, 1983; Senge, 1990, Weick, 1990). Thus, instance-based learning provides the theoretical foundation for ICM, with the analysis of a set of the individual's actual decisions providing the best indication of his/her mental model. That is, given that the instance is the basis of learning how to make the decision and that the storage/retrieval of instances (or chunks of instances) is the mechanism used in applying the individual's knowledge to new decision making situations (Gonzales et al., 2003; Simon & Gobet, 1996), then it is logical that the analysis of multiple decision instances should provide the most direct and rational basis for determining the decision maker's tacit mental model, as well as revealing any related explicit knowledge.

The cognitive science literature also suggests that decisions are based on both explicit (declarative) and tacit (procedural) knowledge (Sun et al., 2001; Sun & Peterson, 1998). Declarative knowledge is explicit, generic knowledge (e.g., instructions, relevant data, written procedures) concerning how to make a decision in a specific domain; once learned, declarative knowledge is mentally accessible to the decision maker and can, therefore, be used to explain how and why a decision was made. On the other hand, procedural knowledge is tacit, experiential knowledge developed by making multiple decisions in a problem domain; it is based on subconscious processing and is mentally inaccessible to the decision maker in explaining how or why a decision was made (Fiol & Huff, 1992; Gonzales et al., 2003; Schraagen, 1993;Tsoukas, 2003;). Both types of knowledge are critical in making decisions in complex environments, and both types of knowledge vary with decision making experience, with tacit knowledge representing at least a significant, if not the dominant, portion of the overall decision making knowledge (Gonzales et al., 2003; Reber, 1989; Sun et al., 2001). However, since tacit knowledge is inaccessible to the decision maker in explaining how/why a decision is made, the interviewing of decision makers to determine cause/effect relationships, decision schemas, or expert rules will omit a major component of knowledge employed by the decision maker, specifically, all or a major portion of the tacit procedural knowledge. On the other hand, ICM analyzes actual decisions, each of which utilizes all the tacit and explicit knowledge the decision maker(s) deems relevant (Sun et al., 2001; Sun & Peterson, 1998).

Theoretical Justification from Artificial Intelligence

In the literature of artificial intelligence (AI), instances (usually referred to as cases) are also a mainstay of learning and decision making in both repetitive and novel situations (Gilboa & Schmeidler, 2000). Case-based reasoning and case-based learning state that learning and reasoning occurs through the storage and analysis of many cases containing the same triplets of key factor values, decisions, and outcomes discussed in cognitive science (Gilboa & Schmeidler, 2000; Kolodner, 1993; Reisbeck & Schank, 1989). This suggests that the decision instance (case) should be the elemental unit of analysis for ICM's mental model discovery.

Additionally, one of the highest priority research topics and the fastest growing areas of AI applications during the past 15 years has been knowledge discovery in databases (KDD). KDD is defined as the nontrivial extraction of implicit, previously unknown and potentially useful information from databases: i.e., the discovery of patterns that are both interesting and sufficiently certain to be of value to the decision makers (Marakas, 2003). The theory and technologies of KDD are directly applicable to the analysis of multiple decision instances in ICM. That is, if we think of the multiple decision instances as multiple tuples in a relational table, and the factors and decisions as the attributes, then KDD theory in general, and self-learning inductive technologies in particular (e.g., GMDH (Farlow, 1984; Ivakhnenko, 1971)), should provide a viable source of relationships between decisions and key factors: i.e., a mathematical representation of the decision maker's mental model.

Instance-Based Cognitive Mining

The sections below discuss the ICM assertions, ICM instances, the ICM process, and the associated software in our initial implementation.

ICM Assertions

ICM asserts, based on the literature discussed above, that decision making integrates explicit declarative knowledge (e.g., facts, data, written procedures, rules) and tacit procedural knowledge (i.e., mental models); an individual's knowledge varies over time and with experience, but tacit knowledge is always a significant, if not the dominant, component. Further, a decision maker usually cannot describe his/her own tacit mental model; rather, a mental model must be hypothesized, and tested by observing actual decisions or actions (Argyris & Schön, 1996, pp.16-17). Within a specific domain, a decision maker relies on certain key factors in evaluating decision alternatives; different values for one or more of these key factors may lead to different decisions. That is, when presented with a decision situation in a specific problem domain, a decision maker first determines the appropriate values of self-selected key factors and then "executes" his/her mental model of the decision domain (incorporating any relevant explicit knowledge) to evaluate the alternatives, estimate potential outcomes, and make the best decision. Finally, ICM asserts that inductive analysis of multiple decisions made by an individual (or small team of individuals) in a specific problem domain can be employed to derive a mathematical representation of the mental model actually used; this derived model will be a function of the individual's key factors and will explicitly relate how the decision maker implicitly uses and weighs these key factors in decision making.

ICM Instances

ICM defines an instance as the values of key factors that together define a specific situation in a decision domain. With respect to the cognitive science definition of "instance", our definition could more appropriately be termed a "pre-decision instance"; after the decision is specified, it could be termed a "post-decision instance"; and after the decision has been implemented and audited, it could be termed an "evaluated instance". To simplify the discussion below, we refer to all three as "instances", and let the context determine the specific type of instance.

ICM Process

Based on the requirements and characteristics dictated by the three literatures discussed in the previous section, we propose the ICM system architecture shown in Figure 1. This logical design, independent of any specific software or hardware implementation, is discussed below in terms of the modules of Figure 1. It follows the ICM multi-step process that begins by delineating the decision domain and generating, in the Key Factor Module, an initial set of potential key factors for making decisions.

Key Factor Module

Potential key factors can be generated by applying any one of several techniques. For example, the decision maker(s) can be asked, after a brief tutorial and explanation, to draw an influence chart, an easily learned diagramming technique that depicts the outcome variables (decisions) that are of interest in the decision domain and the input variables (key factors) that influence those decisions (Clemen, 1991; Powell & Baker, 2007). Influence diagrams provide a method of specifying an initial set of explicit, and perhaps tacit, key factors used in the decision maker's mental model.

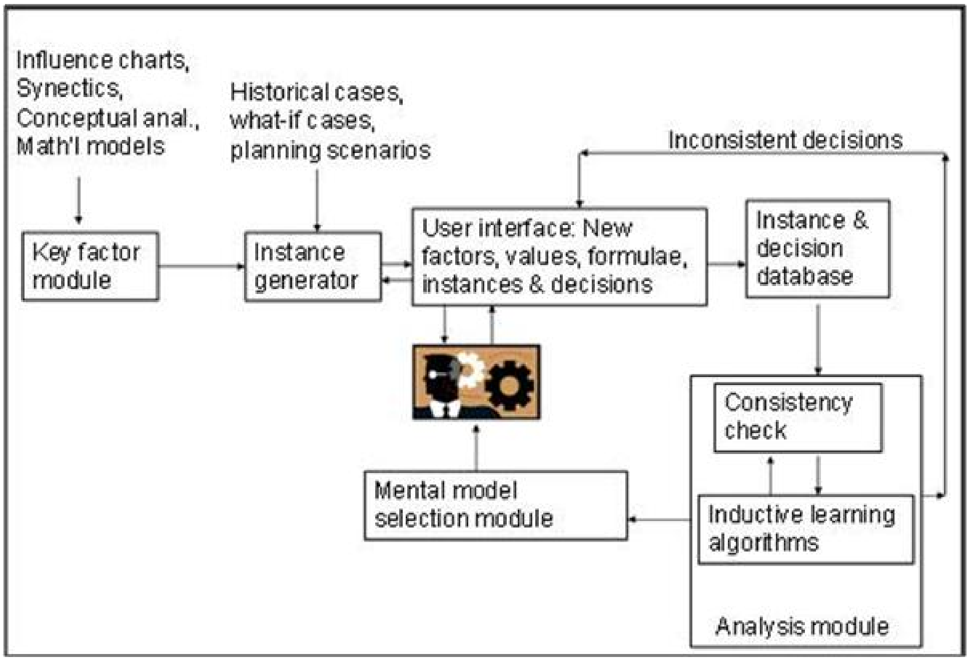


Figure 1: A schematic diagram of Instance-Based Cognitive Mining

Alternatively, structured brainstorming, such as Synectics (Gordon, 1961; Prince, 1970) or its electronic equivalent of group support systems (Nunamaker, Briggs, Mittleman, Vogel, & Balthazard, 1997), can be used in small groups to discover factors that decision makers use to structure and form their decisions. Synectics is both a theory and process that "applies to the integration of diverse individuals into a problem-stating, problem-solving group" (Gordon, 1961, p. 3). Such structured brainstorming is sometimes used in other cognitive mapping techniques such as causal mapping.

Another source of the potential key factors might be a mathematical model of the decision domain. For example, a visual interactive simulation model of a Ford Motor Company engine assembly plant helped to generate key factor values and decision instances that described situations in which machine failures occurred. Such machine failures required human decisions to determine who should be assigned to repair the machine and when the repair should be scheduled (Robinson, Alifantis, Edwards, Ladbrook, & Waller, 2005).

A final source of potential factors might be the conceptual analysis of the decision domain by strategic planners (Hodgkinson, Brown, Maule, Glaister, & Pearman, 1999; Senge, 1990), generating key factors and decision instances (scenarios) representing future states of the industry/world. Such scenarios have long been the cornerstone of strategic planning and changing/enhancing top managers' mental models at Royal Dutch/Shell (Schoemaker, 1993; Senge, 1990).

Instance Generator Module

From this set of potential factors, the researcher builds a set of 15 or more (depending on the number of factors) rationally formulated instances that are representative of the decision space; i.e., a set of instances that optimally covers the decision space with a minimal number of instances (to avoid decision maker fatigue). Such instances, specified in the Instance Generator Module (Figure 1), may include a firm's historical cases, future competitive scenarios proposed by corporate planners (Senge, 1990), and/or the 'what-if' cases generated by a decision maker exploring uncertain decision variables in a mathematical model of the decision domain (Steiger & Sharda, 1996).

User Interface Module

The instances, with each instance containing appropriate values of the key factors but without decisions, are then presented as a set to the decision maker(s) (via the User Interface Module in Figure 1). Since different individuals may rely on different key factors in the same decision domain, the ICM process encourages the decision maker to specify additional key factors not included in the initial instances. If a new factor is independent of the original factors, the user is prompted for a representative range of values for that factor. Alternatively, if the new factor is an explicit, quantifiable function of one or more of the original factors, the user is prompted for a relevant algebraic expression or formula. Finally, if the new factor is tacitly dependent on one or more of the presented factors, the decision maker can intuitively specify a value of the new factor for each instance. These additional factors and the appropriate ranges, intuitive values, or algebraic expressions are then returned to the Instance Generator to generate a new set of instances that includes both the original and the new key factors. Note that by allowing the decision maker to specify new factors that are either tacitly or explicitly dependent on the initial factors, ICM provides for knowledge "chunking", thereby extending the limitations of 7 ± 2 factors normally considered simultaneously in short-term memory (Simon, 1974).

The decision maker may also create one or more additional instances by specifying a value for each of the factors; these instances, which may represent historical situations with which this specific decision maker is familiar, are then included with the original instances. By providing this capability, ICM eliminates the potential problems caused by a researcher's initial set of instances (unalterable by the decision maker) introducing a bias in the decision making and the resulting mental model estimate.

Once all new factors and instances are generated, the user is prompted for a decision in each instance, basing that decision on what s/he considers to be key factors, the relative values of those key factors, and his/her mental model. Note that in this process, the decision maker uses his/her mental model to implicitly or explicitly select any subset or superset of the factors initially presented in each instance. For example, if values for factors A, B, C, and D are initially included in each instance, and the decision maker thinks that only factors A and C along with the new factor E are important, then s/he quite naturally specifies the decisions based on only the values of A, C and E, ignoring B and D values completely.

Analysis Module

The resulting instances, along with the associated decision for each instance, are stored in the Instance and Decision Database (Figure 1). The set of instances is then fed into the ICM Analysis Module for processing. This module features both linear regression software (for linear estimates of mental models) and inductive AI learning algorithms (for nonlinear estimates of mental models). One such nonlinear inductive algorithm is the Group Method of Data Handling (GMDH) (Farlow, 1984; Ivakhnenko, 1971), a family of inductive self-organizing algorithms that can be thought of as a cross between (non)linear regression and artificial neural networks. Each algorithm in the Analysis Module generates one mental model estimate based on the decision instances.

As part of the ICM analysis, the decisions are checked for inconsistencies. Note that these are inconsistencies with respect to one set of decisions; i.e., each set of decision instances presented to the decision maker (or team) is analyzed as an single entity, with the analysis and inconsistencies totally independent of the other sets of decisions made by other decision makers/teams, and independent of any theoretical model. Inconsistent decisions may be caused by clerical error, lack of information/knowledge, lack of concentration, lack of real world consistency, decision maker fatigue, and/or inadequate mental model structure (Forman & Selly, 2001). Inconsistencies may not be necessarily bad; in fact, they may be the source of generating new knowledge by the decision maker as s/he searches tacit and/or explicit knowledge to determine the cause of the inconsistencies, whether it be heretofore unrealized key factors, factor weighting, etc. (Gabbay & Hunter, 1991)

Inconsistent decisions are revealed by implementing a version of Wagner’s (1995) ‘all save one’ algorithm in which the ICM Analysis Module is repeatedly called with *(n-1)* instances, excluding a different instance (row) in each different run, and comparing the results; a result with a significantly higher value for the coefficient of determination, adjusted R^2, and/or a more parsimonious model indicates that the inconsistent decision has been excluded. Any inconsistent decision(s) are fed back to the decision maker, including comparisons with other instances to highlight the inconsistency; the decision maker/team is then requested to either validate the decision or correct the inconsistency.

Model Selection Module

ICM’s Mental Model Selection Module is used to select the “best” model or sub-model of the several mental model estimates generated by the different algorithms in the Analysis Module. The selection criterion is based on a combination of explanatory power (adjusted R^2, and model parsimony; other evaluation criteria that could be included in the selection process are sufficiency (whether the estimated model is sufficient to depict the mental model), necessity (whether all terms in the estimated mental model are required to depict or understand the mental model), etc. (Steiger, 1998).

ICM’s best estimate of the mental model is then displayed to the decision maker for evaluation.

Software Support of the ICM Process

The software support of our ICM implementation focuses on the Instance Generator, User Interface, Instance Database, and Analysis modules. That is, the initial set of decision instances is input by the researcher via interactive PHP forms; PHP is a scripting language that is especially suited for developing web-based, interactive forms (Ullman, 2001). Interactive PHP forms (e.g., Figure 2) are also used to capture additional key factors specified by the decision maker and the corresponding ranges of values (via clicking the “AddFactor(1)” option in Figure 2), formulae (via clicking the “AddFactor(2)” option), or intuitive values (via clicking the “AddFactor(3)” option). Interactive PHP forms are also used to add instances (via clicking the Add Cases option) and to specify decisions (via keying in the decision into the “Results” column) for each instance. The instances are stored in a MySQL database (Ullman, 2001).

The Analysis Module has been implemented as a loosely-coupled, multiple algorithm software system, including Microsoft Excel with PHStat addin (Stephen, 2007) (to generate linear estimates of the decision maker's mental model), inductive learning networks (to generate nonlinear estimates of the mental model), and an implementation of Wagner's (1995) 'all-save-one' algorithm (to detect a decision in one instance that is inconsistent with decisions made by the same individual/team in other instances). Inductive learning networks represent a family of self-organizing inductive learning algorithms that employ multi-layered, cascading networks of inter-connected nodes used to model nonlinear relationships. Inductive learning networks generate a mathematical model that, in this application, explains variations in the instance decisions based on variations in the values of the corresponding key factors in each instance. Since inductive learning networks are a family of algorithms, with each member of the family based on slightly different modeling assumptions (e.g., partial quadratics, partial cubics, ratios of polynomials, etc.), two such algorithms are used in our ICM prototype implementation: KnowledgeMiner, available at www.knowledgeminer.com/, and PolyAnalyst's FindLaws, available at www.megaputer.com/.

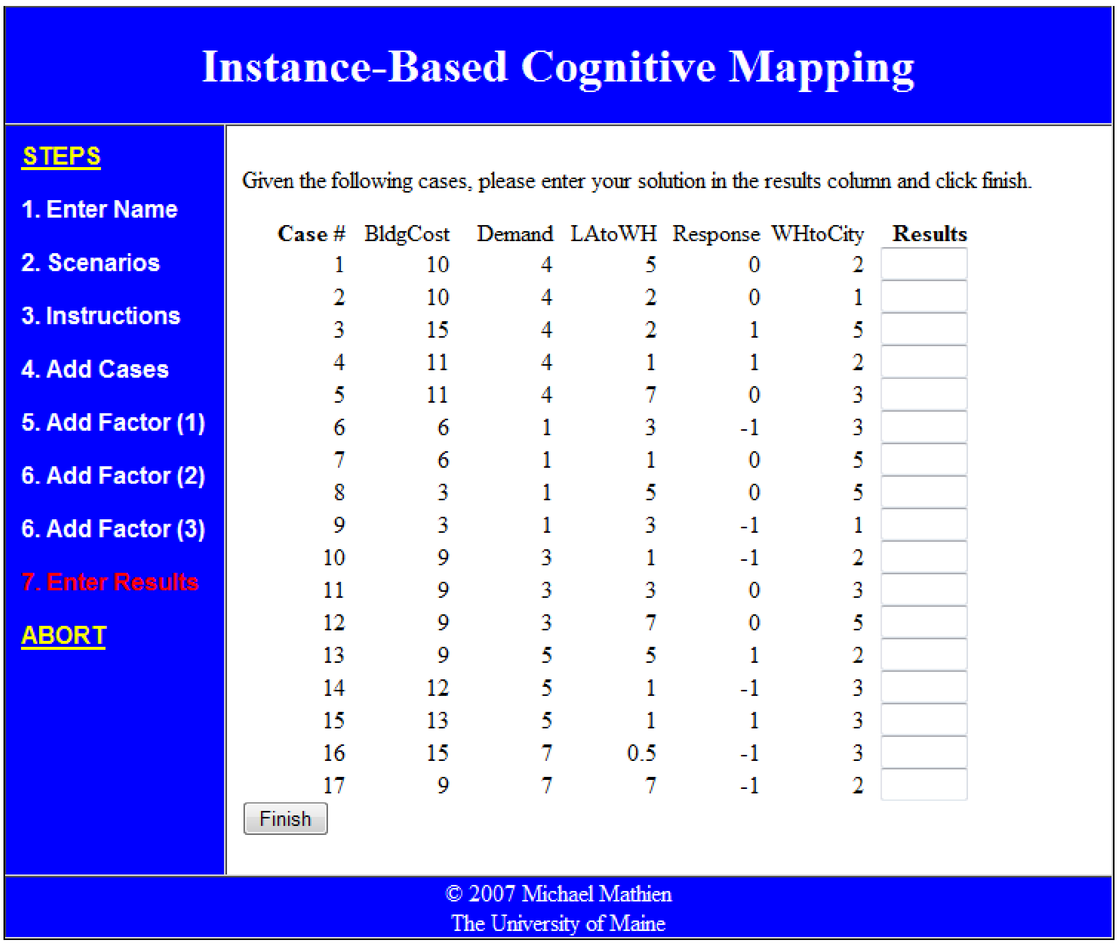


Figure 2: Screenshot of ICM--instance key factor values with column for decisions (labeled "Results").

KnowledgeMiner (Müller & Lemke, 2003) is a commercially available implementation of the GMDH algorithm that provides (non)linear additive (quadratic) polynomials to estimate the decision maker's mental model, without requiring any prior knowledge concerning which key factors to include, or the form of any relationships between those key factors (i.e., interactions terms, squared terms, etc.). For each model, KnowledgeMiner also calculates and provides various predictive statistics (PESS, MAPE, R^2, adjusted R^2, and descriptive power) for the estimated model, as well as a chart of the estimated vs. actual values of the dependent variable. In this application, adjusted R^2 provides a measure of the consistency of the indicated decisions.

PolyAnalyst's FindLaws is another commercially available implementation of inductive learning networks, but instead of an additive quadratic polynomial model, it produces a nonlinear estimate of the decision maker's mental model based on a ratio of polynomials. Like GMDH algorithms, it requires no prior knowledge concerning which key factors to include or the form of any relationships between those key factors (i.e., interactions terms, squared terms, etc.). For the final model, PolyAnalyst FindLaws also provides the standard error, standard deviation, model significance, and coefficient of determination, as well as a chart of the estimated vs. actual values of the dependent variable.

Since neither of these software packages was available to us as a dynamically linked library routine (callable from a PHP form), our initial prototype implementation of ICM requires manual intervention in the Analysis Module; future versions will fully automate this module.

A Preliminary Test of ICM

As a preliminary test of ICM in a business environment, assume the decision domain is that of the classical warehouse location problem (Geoffrion, 1976) in a new market area that includes thirteen cities, all located in Central Texas, with a single source of supply in Los Angeles. The problem is to determine the best number of warehouses required to serve the marketing cities at the lowest overall cost. There must be at least one, and no more than thirteen, warehouses, with each selected warehouse located in one of the demand cities. Each warehouse must have sufficient throughput capacity to serve the total Central Texas market area.

To illustrate the ICM analysis characteristics and capabilities using small teams of student decision makers, we integrated the warehouse location problem into a (required) MBA Operations Analysis course. Of the 30 MBA students participating in this study, the average age was 30, with a third of the participants being women. Further, one third of the participants had managerial experience (on average, 4.2 years of managerial experience). Undergraduate degrees held by our subjects included 40% with technical degrees, 33% with business degrees, and 27% with other degrees.

The class was divided into thirteen small teams, and each team, in succeeding weekly homework assignments, developed an influence diagram of the warehouse location problem, completed spreadsheet sensitivity exercises using the events (key factors) noted in their influence diagrams, and developed a linear programming approximation of the problem. This course integration was aimed at developing and enriching the students' expertise (i.e., their mental models) in this decision domain.

The influence diagrams, which help externalize the tacit and explicit knowledge in each team's mental model, were analyzed to generate an initial set of four factors that were common to all thirteen of the teams' influence diagrams. These four factors consisted of the forecasted product demand, d, the fixed building costs, f, of each warehouse, the unit transportation cost, T, from the Los Angeles factory to the warehouses in Central Texas, and the unit transportation costs, t, from the warehouse to the demand city.

After the linear programming model was developed by each team, we introduced one additional factor, the delivery service s, a categorical variable indicating the time between placing an order and receiving the ordered goods from the closest warehouse (1 = same day delivery, 0 = next day, and -1 = two business days delivery).

This set of five factors, representing an initial factor set based on the tacit and explicit knowledge of each team (as indicated by their influence diagrams), was then used to generate a set of 17 decision instances that efficiently covered the decision space (Figure 2). Note that there is an inherent conflict in the number of decision instances used in the ICM process; i.e., more instances would provide greater statistical reliability and confidence in the ICM estimate of the mental model. However, more instances would also increase the risk of exceeding the decision maker's attention span, introducing decision maker fatigue and decision inconsistencies, resulting in lower statistical reliability in the ICM estimate. In preliminary tests of the ICM software, we included as many as 48 instances in the warehouse model application (using the same five factors) and found that the ICM process produced the same results regardless of the number of instance down to the 12-15 instance level, as long as the remaining instances covered the decision space relatively well. Unfortunately, the 48 instances required well over an hour to complete by the decision makers, leading to decision maker fatigue and explaining some of the inconsistency in decisions.

The five factors in this set of instances were also tested for colinearity by calculating the variance inflationary factors (Levine, Stephan, Krehbiel & Berenson, 2005); based on the criteria developed by Snee (1973) there was little evidence of colinearity.

Each team was presented with the 17 different decision instances, consisting of the five factors and a blank input column for decisions (Figure 2). Decision makers were encouraged to ignore any irrelevant factors, to add any new instances with which they were familiar by clicking the "Add Cases" button, and add new key factors as needed by clicking one of the three buttons labeled "Add Factor" (1), (2), or (3). They were then asked to analyze the decision instances, determine the best number of warehouses, *n*, to build in each instance (being as consistent as possible), and specify the team's decision in the ICM "Results" column. Note that each team was required to reach a consensus decision on each decision instance; this resulted in significant discussion among team members as they detected, evaluated, and explained their individual differences in decisions based on (the externalization of) their individual assumptions and mental models.

Each team also filled out a questionnaire requesting additional information concerning their mental model of the decision domain, including a self-estimate of their decision consistency, a list of factors that they used in the decision making process, a list of those factors explicitly ignored, and a mathematical representation of their mental model, if any.

Expected Results

Geoffrion (1976) developed a mixed integer/linear programming model of the warehouse location problem during a consulting assignment at Hunt-Wesson Foods. Using this mathematical model, several simplifying assumptions (specifically, that the forecasted demand was the same for each city, the fixed building costs of each warehouse were the same at each potential warehouse location, and all warehouse locations were in the same shipping zone from the product source), human expertise, and mathematical manipulations, he developed an insight generating simplified auxiliary model for the optimal number of warehouses, *n*, in an area of *A* square miles as:

$$n=(A/3.05)*(d*t/f)^{2/3}, \qquad \text{Eq. 1}$$

where d = the demand in each city, t = the unit transportation cost from the nearest warehouse to the demand city, and f = the fixed costs associated with building each warehouse. Geoffrion's analysis did not include the irrelevant (to the optimized results) factory-to-warehouse transportation cost, T, or the delivery service factor, s, in the model.

In a pre-test application of the ICM process to the warehouse location problem given to different students in a previous semester, we found that the student decision makers reported using additive formulae in their mental models, with positive coefficients associated with demand, transportation costs and delivery service (i.e., increases in the values of these key factors from instance to instance resulted in increases in the optimal number of warehouses), and a negative coefficient associated with the fixed warehouse building costs (i.e., increases in the fixed building costs resulted in decreases in the optimal number of warehouses). These straightforward additive mental models were to be expected given the students' relative inexperience in the decision domain.

To provide a component of diversity in the results generated by the 13 teams in the current test, we emailed suggestions to just over half of the teams (specifically Team #1 through #7) that the optimal number of warehouses is best determined by using a multiplicative relationship between the key factors versus an additive relationship. That is, we coached these teams to "multiply key factors instead of adding them, and divide key factors instead of subtracting them." Note that if the decision maker's mental model is totally tacit, this coaching should have little or no effect.

Experimental Research Issues

We tested two research hypotheses. Stating these hypotheses in terms of mediated models (Baron & Kenny, 1986), the cognitive science literature suggests that there is a mediating causal process (a mental model in this research), *MM*, that provides the mechanism through which the factors describing decision making instances, *F*, affect the actual decisions, *D*, made in the instance (Figure 3). The artificial intelligence literature suggests that inductive learning networks, (e.g., GMDH), can be applied within the ICM process to generate a mathematical representation of this mental model. Thus, our first hypothesis is:

> H1: The ICM-derived mental model representation, *DMM*, provides the mediating causal process through which the set of factor values, *F*, affects the actual decisions, *D*, made by the team.

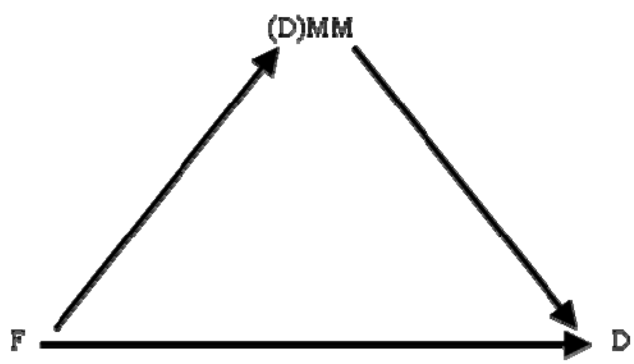


Figure 3: Mediated model with the mental model as the mediating causal process.

Note that the set of factor values, *F*, (that describe each decision situation) may contain a subset of the initial factors presented to the decision maker(s) or a superset of those factors after inclusion of additional factors (and associated values) specified by the individual/team. Further note that the ICM process requires no *a priori* knowledge of which factors the team actually used in its decisions, the algebraic relations between these key factors (e.g., cross terms, squared terms, inverse relationships, etc.), or the relative weighting of the factors or terms.

The cognitive science literature also suggests that a decision maker's mental model consists of both explicit and tacit knowledge, with tacit knowledge representing at least a significant portion, and often a dominant portion, of the overall knowledge. Explicit knowledge is accessible to the decision maker in explaining how/why a decision is made, i.e., in providing a self-reported mental model. A self-reported mental model may be expressed as an algebraic relationship, a directional relationship (e.g., "as *x* increases, *y* increases"), or an (unspecified) tacit relationship with only the key factors reported. Thus, the first part of our second hypothesis is:

> H2a): The ICM-derived mental model representation is consistent with each team's self-reported algebraic, directional, or tacit mental model relationships, and each team's key factors.

Finally, the literature suggests that tacit knowledge is often inaccessible to the decision maker in explaining how or why a decision was made; i.e., in providing a self-reported mental model. That is, a self-reported relationship may be inconsistent with the actual decisions made by the individual or team, due to this inaccessibility of tacit knowledge or a desire for political correctness (e.g., not wanting to admit to racial bias), etc.(Argyris & Schön, 1996). Such inconsistencies can be detected in the ICM process and form the basis for the second part of our second hypothesis:

> H2b): Any significant differences between the ICM-derived mental model representation and the team's self-reported mental model or key factors are not consistent with the team's actual decisions.

Actual Results

After developing (via the influence diagrams), modifying (based on factors and instances added by the team), and saving (in the MySQL database) the instances and the team's final decisions, the set of instances (including decisions) was passed to the Analysis Module. In the Analysis Module, ICM generated three potential estimates of each team's mental model: 1) a linear estimate based on multiple linear regression, using the best-subsets approach to model building as suggested by Levine et al. (2005, pp. 637-642); 2) a nonlinear estimate based on the ratio of polynomials algorithm of PolyAnalyst FindLaws; and 3) a nonlinear estimate based on the additive polynomial GMDH algorithm of KnowledgeMiner. The best one of these three mental model estimates was then selected based on model simplicity and adjusted R^2. Adjusted R^2, versus simply R^2, is used since it adjusts for both the number of key factors (i.e., explanatory variables) and the number of decision instances (sample size) used in ICM's estimated model (Levine et al., 2005, p. 580).

Table 1, Column 2 summarizes ICM's best mental model estimate for each of the thirteen teams participating in the test; each estimate is significant at the 1% level indicating that there is a very low probability that we would have obtained the ICM model if the true values of the model coefficient(s) were zero. As seen in the table, ICM discovered additive linear models for seven of the thirteen teams, all of which included negative coefficients for the fixed building costs, *f*, and positive coefficients for all other key factors that were included in the team's model; the lone exception was one negative coefficient for demand, *d*, in Team #13's estimated model. The adjusted coefficients of determination, adjusted R^2, of the seven linear models ranged from 81% to 89% (Table 1, Column 3). For the remaining six teams, ICM discovered nonlinear models with adjusted R^2 values ranging from 82% to 100%.

Table 1. Comparison of ICM Best Model with Team Reported Information

Team	ICM Best Model[#]	Model R^2_{adj}	Model Source*	Inconsistent Instance	Questionnaire Information: Self-Reported Mental Model Formula	Questionnaire Information: Key Factors[#]
1	n=4.6-.49f+.66d+1.1s+1.8t	87	XLPH	#10	cost=$nf(c^s)+(dt)$× distance	f, s, t, d
2	n=3.8-.27f+.36d+.17T+.7s+.7t	83	XLPH	#16	n directly proportional to d, T, & t; n inversely proportional to f	f, d, t, T
3	n=7.6$(dt)/f$	99	PAFL	–	n=(3/4)(t/100)(d/f)	t, f, d
4	n=.44-.35f+1.6d+.75s+1.2t	87	XLPH	–	none ("subjective")	d, t, s, f
5	n=15$td/(Tf)$	91	PAFL	–	n=(.4t)(.1d)(.1s)/{(.35f)(.05T)}	t, f, d, s, T
6	n=4.94$dt/(f$-2.2$s)$	98	PAFL	–	n=$(td/f)+s$	f, d, t, s
7	n=5.23$dt/(f$-1.87$s)$	98	PAFL	–	n=(.3d)(.4t)/(.3f), adjusted for s	f, d, t, s
8	n=6.5dt/f	82	PAFL	#17	n directly proportional to d, T, & t; n inversely proportional to f	d, t, f
9	n=4.9-.82f+.81d+2.0t	83	XLPH	#9	none ("subjective")	f, t, d
10	n=5.1-.39f+.74d+.96s+.66t	81	XLPH	–	n directly proportional to d, T, & t; n inversely proportional to f	t, d, f, s
11	n=4.9-.45f+.48d+.27T+1.4t	85	XLPH	#3	none ("subjective")	f, d, t
12	n=7-f+dt	100	KMNL	–	n=dt-f-constant	d, t, f
13	n=14.3-.67f-.29d	89	XLPH	–	none ("subjective")	t, d, f, s

[#] n=best number of warehouses, d=demand in each city, t=unit transportation cost from the nearest warehouse to the demand city, f=fixed costs associated with building each warehouse, T=factory-to-warehouse transportation cost, and s=delivery service factor.

* Model Source: PAFL (PolyAnalyst FindLaws); XLPH (Microsoft Excel, PHStat addin); KMNL (Knowledge Miner NonLinear with Stepwise Regression)

Note that the seven linear ICM mental model estimates (Team #1, #2, #4, #9, #10, #11, and #13) are models that are linear in the original factors (i.e., demand in each city, *d*, unit transportation cost from the nearest warehouse to the demand city, *t*, fixed costs associated with building each warehouse, *f*, factory-to-warehouse transportation cost, *T*, and delivery service factor, *s*) and, therefore, linear regression is appropriate for finding the best fit coefficients for these estimated mental models. One ICM mental model estimate (#12) contains a cross product of the original factors (i.e., demand times warehouse-to-city transportation cost), and, according to Seber & Wild (2003, p. 5), since it is linear in the parameters (coefficients) of its terms, linear regression is also appropriate for finding the best coefficients for this decision team's mental model estimate. The other five ICM mental model estimates (Team #3, #5, #6, #7, and #8) are provided by the proprietary algorithm of PolyAnalyst FindLaws and, whereas they are nonlinear with respect to the original factors, the proprietary algorithm's term when considered as a whole is linear in the one best fit parameter (coefficient) for the single overall term (e.g., the 5.23 coefficient in team #7's estimated mental model). Thus, linear regression analysis is appropriate for all thirteen mental model estimates, since "the important requirement (for linear regression) is that the expression should be linear in the parameters (i.e., coefficients)" (Seber & Wild, 2003, p. 5). Likewise, the goodness-of-fit measures, adjusted R^2 and the significance *F*-test are appropriate for all thirteen models.

Four of the nonlinear models closely followed the insight generating ratio of variable costs to fixed cost shown in Eq. 1. All three of ICM's discovery algorithms (Excel with PHStat, KnowledgeMiner, and PolyAnalyst) (Table 1, Column 4) were used in this application.

These results confirm the first research hypothesis; i.e., H1: the ICM-derived mental model representation provides the mediating causal process through which the set of key factor values affects the actual decisions made by a team, with no *a priori* knowledge of which factors the team actually used in its decisions, the algebraic relations between these key factors (e.g., cross terms, squared terms, inverse relationships, etc.), or the relative weighting of the factors. Further, ICM provided two goodness-of-fit measures (adjusted R^2 and the *F*-test) indicating that the mental model estimate generated by ICM was, in fact, statistically valid.

Three teams utilized the ICM option of adding a key factor; in fact, one team added several new factors. All added factors were functions of existing key factors; i.e., the teams that added factor(s) demonstrated knowledge "chunking" (Simon, 1974). In addition, one team tried to add several new decision instances, another ICM option; unfortunately, they were thwarted when they exceeded the maximum number of instances (20) allowed by ICM at the time.

In a breakdown of the self-reported information from the questionnaires (Table 1, Columns 6 & 7), five of the thirteen teams reported using an algebraic relationship, three others (Team #2, #8, and #10) indicated directional relationships, four teams (Team #4, #9, #11, and #13) indicated a tacit relationship (i.e., "subjective" or "none") for the mental model used, and one team (Team #1) listed a cost formula as their mental model (something to minimize?). Note that the self-reported relationships shown in Table 4, Column 6 reflect those written on the team's questionnaire response; some teams obviously adjusted their relationship to generate an appropriate number of warehouses.

Of the five teams that reported an algebraic relationship, ICM's model matched the team's reported relationship (except for omitting *s* in the reported model for Team #5). Of the three teams that reported directional relationships (Team #2, #8, and #10), ICM generated a model that was consistent with the reported relationships for each of the three teams. Of the five teams that indicated a tacit relationship (Team #1, #4, #9, #11, and #13), ICM's best model used the same key factors for three teams (Team #1, #4 and #9), included one additional factor (*T; P*-value significant at the 5% level*)* for one of the teams (Team #11), and omitted two factors (*s* and *t*) for the fifth team (Team #13; the P-values for both s and t were statistically insignificant at the 10% level).

These results confirm the second research hypothesis; H2a: the ICM-derived mental model representation is consistent with the team's self-reported algebraic, directional, or tacit relationships, and the team's self-reported key factor; or H2b: any significant differences between the ICM-derived mental model representation and the team's self-reported relationships or key factors are not consistent with the team's actual decisions.

Inconsistencies Highlighted by ICM

The purpose of ICM is to discover a mathematical representation of the decision maker's (or decision team's) mental model. However, since ICM is based on actual decisions, it also provides feedback to each team concerning possible inconsistencies inherent in their decisions and/or inconsistencies between their decisions and their self-reported mental model characteristics. Knowledge of such inconsistencies, along with ICM's estimated mental model, may promote improvements in the individual's mental model (i.e., it enhances decision maker learning). That is, discrepancies between the externalized mental model and current (inconsistent) decisions stand out and can be analyzed because comparison with the mental model is made possible (Weick, 1990). Further, discussions of such inconsistencies by individuals within the team represent another form of tacit knowledge externalization, as suggested by Nonaka (Nonaka & Takeuchi, 1995). Note that these are inconsistencies within the decisions of a single team and/or the reported information of that team; i.e., each team's set of decision instances is analyzed as an independent entity, with the analysis and inconsistencies totally independent of the other individuals/teams, and independent of any theoretical model (e.g., Eq. 1).

ICM detected three types of inconsistencies in this test. Inconsistencies-of-application are instances detected by the "all-save-one" algorithm that, when eliminated, cause the adjusted R^2 of ICM's estimated model to increase significantly. Such inconsistencies might reflect clerical errors when entering decisions, lack of concentration or mental fatigue on the part of the decision makers, mental model differences between members within the team, etc. (Forman & Selly, 2001). Five of 13 teams experienced inconsistencies-of-application (see Table 1, Col. 5); e.g., instance #10 for Team #1, when omitted, increased adjusted R^2 from 74% to 87%. These inconsistencies are fed back to the team, along with ICM's best mental model estimate, for reconsideration and (possible) correction. Inconsistencies-of-reporting are indicated by a key factor being reportedly ignored by the team on its questionnaire, but included in ICM's estimated model; this type of inconsistency may demonstrate either the team's tacit use of a factor or inconsistencies-of-application. Two teams demonstrated an inconsistency-of-reporting (Team #2 and #11); e.g., Team #2 reportedly ignored the service factor, s, but the P-value for this factor, when added, was significant at the 5% level, and including it in the ICM model raised the adjusted R^2 from 68% to 83%. Finally, inconsistencies-of-weighting are indicated by a key factor being listed in the team-reported key factors as used in the decisions, but excluded in the ICM estimated model; i.e., the team placed insignificant weight on the reported key factor for it to matter. This may indicate a discrepancy between the self-reported model and their actual mental model (Argyris & Schön, 1996). Two teams (#5 and #13) demonstrated inconsistency-of-weighting; e.g., Team #13 reported using the service factor, s, and the transportation, t, in their decisions, but ICM's estimated model excluded these two factors and calculated an adjusted R^2 with the remaining factors of 89% and an associated F-test which is significant at the 1% level. Both factors were statistically insignificant (based on P-values insignificant at the 10% level) when included in the model.

Analysis of Team Results

As for specific team analysis, Team #6 was very consistent in their decisions as indicated by the adjusted R^2 of 98%. PolyAnalyst generated ICM's model estimate as $n = 4.9 * d * t / (f - 2.2 * s)$. This fits well with the self-reported mental model provided in Team #6's questionnaire: $n = (t * d / f) + s$; i.e., if $s = +1$ (indicating same day service) you would expect that additional warehouses might be needed due to the relatively rapid response requirement, a condition that is met by both the team's formula and the PolyAnalyst model. Likewise, a Service value of -1 (two business days) would indicate that fewer warehouses might be needed to meet this requirement, a condition that is again met by both models. Excel/PHStat generated a linear model with a significantly lower adjusted R^2 (82% versus 98%) for this team's mental model, whereas KnowledgeMiner's nonlinear model was very complicated (including polynomial terms raised to the fourth power) and no improvement in adjusted R^2.

These results highlight several important characteristics of inductive learning networks. First, categorical variables such as the service level can be included in the ICM instances and, if used by the decision maker, are logically incorporated by the inductive learning networks into its estimate of the mental model. In addition, factors that are considered irrelevant by the decision maker are routinely discarded and appropriately omitted by the inductive learning networks; e.g., the FindLaws model above suggests that the factory-to-warehouse transportation rate, T, was irrelevant to Team #6's decisions, a finding that was consistent with key factors listed in this team's questionnaire and, also, consistent with reality since all thirteen warehouses are in the same shipping zone.

Team #4 was less consistent in their decision making, as indicated by an adjusted R^2 of 87% with the following model generated by ICM (using linear regression plus PHStat's Best Model algorithm): $n=-.44-.35f+1.6d +.75s+1.2t$. This additive model is reasonable in that higher values of demand, d, transportation costs, t, and/or service level, s, lead to more required warehouses, whereas higher warehouse building costs, f, lead to fewer required warehouses (i.e., a negative coefficient). The nonlinear model estimated by PolyAnalyst provided a higher adjusted R^2 but was complicated by multiple imbedded if-statements, whereas the nonlinear model produced by Knowledge Miner included high order polynomial terms with no improvement in the adjusted R^2. Note that ICM's best model for this team is dimensionally inconsistent; i.e., warehouse building costs ($) are inconsistent with transportation costs ($/unit) and demand (units), and all terms are dimensionally inconstant with the (unit less) number of warehouses. This provides the basis for feedback to the decision maker suggesting that a multiplicative model (versus an additive model) might provide dimensional consistency and improve decision accuracy in the domain.

Summary and Conclusions

The purpose of this research was to propose a theoretical justification for, and describe an implementation of, instance-based cognitive mining (ICM), a process that analyzes multiple decisions made by an individual (or small team of individuals) in a specific decision making environment, using inductive learning algorithms and regression to generate a mathematical representation of the decision maker's mental model, *explicitly* relating how the decision maker *implicitly* selects and weighs key factors in making decisions. We based our theoretical justification of ICM on three distinct literatures (knowledge creation, cognitive science, and artificial intelligence). Further, we proposed an architecture that is consistent with the implementation characteristics suggested in our literature review and integrates several inductive artificial intelligence technologies, summarizing that architecture in Figure 1. Finally, we developed a prototype implementation of ICM and conducted a preliminary test of the process using MBA students; the results, summarized in Table 1, supported our two research issues. That is, the ICM-derived mental model representation provides the mediating causal process through which the set of key factor values affects the actual decisions made by a team, is independent of researcher bias, and provides two goodness-of-fit measures (significance F-test and adjusted R^2) of the resulting mental model estimate; a lack of structure or fit would be indicated by a low value of adjusted R^2, a significance F-test greater than 0.05, or an overly complicated ICM model. In addition, the preliminary test results supported our second hypothesis that the ICM-derived mental model representation is either consistent with the team's self-reported algebraic, directional, or tacit relationships, and the team's self-reported key factors; or that any significant differences between the ICM-derived mental model representation and the team's self-reported relationships or key factors are not consistent with the team's actual decisions.

ICM provides two unique characteristics that researchers (e.g., Chi, 2007; Gary & Wood, 2007; Weick, 1990) and practitioners (e.g., Schoemaker, 1993; Senge, 1990), have stated promote conceptual change, enhance decision maker learning, and improve corresponding decisions. First, ICM is based directly on the decisions that an individual/team makes (as suggested by Argyris & Schön, 1996); our theoretical justification of ICM strongly suggests that the analysis of multiple decision instances should provide the most direct and rational basis for discovering the decision maker's tacit mental model of the decision domain. Second, the ICM process includes the capability to detect (based on actual decisions) three types of decision inconsistencies: inconsistencies-of-applications, inconsistencies-of-reporting, and inconsistencies-of-weighting (see Inconsistencies Highlighted by ICM section above).

The ICM process is designed to discover and represent relationships between decisions and the key factors (as specified, modified and used by the decision maker). These relationships are represented in the form of parsimonious polynomials, both linear and nonlinear, both additive and ratio-based, providing a wide range of modeling flexibility to capture both tacit and explicit relationships in the decision domain. The ICM process has two primary limitations. One limitation is caused by the number of key factors required in the decision domain; i.e., having fifteen or twenty key factors would require an unwieldy number of decision instances (too many for an executive to consider in a reasonable amount of time), whereas seven key factors could be represented in 30-35 instances (carefully crafted to cover the decision space), and nine factors in 60-70 instances, perhaps an upper bound for a decision maker's attention span. "Chunking" of key factors (Simon, 1974) provides some natural relief as the number of factors increases; e.g., three of our decision teams utilized chunking to simplify the five-variable warehouse location problem in this test. However, we need to explore whether a multi-phased application of ICM might provide additional support when more than nine factors are required.

The second limitation of ICM concerns decision inconsistency between instances; e.g., comparable instances in which an increase in factor, x, causes an increase in decision, d, in one instance, but a decrease in decision, d, in a different instance, all other things equal. Decision inconsistencies result in overly complicated (and probably erroneous) mental model estimates.

The generalizability of this study should be considered in the light of these limitations of the ICM process and the (relatively low) experience of our (MBA) decision makers. According to the Theoretical Justification section of this paper, experienced decision makers can be characterized as developing and using mental models that incorporate higher levels of complexity using "chunks" of key factors and instances, reflecting more tacit knowledge than explicit knowledge, leading to faster and more consistent decision making. Further, the differences between the self-reported (espoused) mental model and the mental model actually used (theory-in-use) will be greater with increased experience as the decision maker justifies his/her decisions in terms of socially, politically, and industrially acceptable standards.

Mapping these characteristics against the advantages and limitations of ICM we suggest that ICM is well designed for experienced decision makers in that: 1) instead of an analysis based on surrogate decision measures or interview data (that omits much of the tacit knowledge), ICM analyzes actual decisions that include all the tacit and explicit knowledge the individual uses in making each decision, leading to better ICM mental model estimates (when compared to our inexperienced subjects); 2) experienced decision makers may take into consideration additional factors, but these factors are often combined (tacitly and/or explicitly) with other factors into “chunks” (specified by the decision maker within the ICM process) to reduce (or at least, not increase) the number of factors and instances presented by ICM; 3) faster decision making leads to more instances that can be presented without exceeding the attention span of the decision maker; 4) increased mental model complexity (i.e., cross terms, squared terms, ratio of terms, and even feedback factors if specified by the decision maker) can be discovered by ICM’s inductive learning networks; and 5) the differences between the espoused and actual mental models of experienced decision makers can be detected and expressed as a by-product of the ICM process.

The preliminary test of our ICM prototype using a seminal decision domain suggests that the ICM process described in this paper is worthy of further development and investigation. However, additional empirical research is needed to explore its application in other decision domains using larger samples of real-world decision makers, instead of the small sample of teams of student decision makers used in our preliminary test.

Acknowledgments

This research was funded, in part, by the Maine Economic Initiative Fund. We wish to thank our graduate students, Josh Wallace and Michael Mathien, for coding and testing the ICM software, and our colleague, Dr. Jason Harkins, for his comments on our empirical study. We also wish to thank the anonymous referees for significantly improving the form and content of this paper. A precursor of this research appeared in Steiger & Steiger (2007).

References

Argyris, C., & Schön, D.A. (1996). *Organizational learning II: Theory, methods, and practice*. New York: Addison-Wesley.

Baron, R., & Kenny, D. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Becerra-Fernandez, I., & Sabherwal, R. (2001). Organizational knowledge management: A contingency perspective. *Journal of Management Information Systems*, *18*(1), 23 – 55.

Carley, K. M., & Palmquist, M. (1992). Extracting, representing, and analyzing mental models. *Social Forces*, *70*(3), 601-636.

Chi, M. T. H. (2007). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Fosniadou (Ed.), *Handbook of research on conceptual change*. Hillsdale, NJ: Lawrence Erlbaum.

Choo, W. C. (1998). *The knowing organization*. NY: Oxford University Press.

Clemen, R.T. (1991). *Making hard decisions*. Boston: PWS Kent.

Davenport, T., & Prusak, L. (1998). *Working knowledge*. Boston, MA: Harvard Business School.

Farlow, S. J. (Ed.). (1984). *Self-organizing methods in modeling: GMDH type algorithms*. NY: Marcel Dekker.

Fiol, C. M., & Huff, A. S. (1992). Maps for managers. Where are we? Where do we go from here? *Journal of Management Studies*, *29*, 267-289.

Forman, E. H., & Selly, M. A. (2001). *Decision by objectives*. New York: World Scientific Press.

Gary, M. S. & Wood, R. E. (2008). *Mental models, decision making and performance in complex tasks*. Retrieved July, 2008 from http://ssrn.com/abstract=1095665

Gabbay, D. & Hunter, A. (1991). Making inconsistency respectable: A logical framework for inconsistency in reasoning, Part I--A position paper. *Proceedings of Fundamentals of Artificial Intelligence Research Conference '91*, 19-32.

Geoffrion, A. M. (1976). The purpose of mathematical programming is insight, not numbers. *Interfaces*, *7*(1), 81-92.

Gilboa, I., & Schmeidler, D. (2000). Case-based knowledge and induction. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, *30*(2), 85-95.

Gonzales, C., Lerch, J., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, *27*, 591-635.

Gordon, W. J. J. (1961). *Synectics*. New York: Harper & Row.

Hodgkinson, G. P. (1997). The cognitive analysis of competitive structures: A review and critique. *Human Relations, 50*, 625, 654.

Hodgkinson, G., Bown, N. J., Maule, A. J., Glaister, K. W. & Pearman, A. D. (1999). Breaking the frame: An analysis of strategic cognition and decision making under uncertainty. *Strategic Management Journal*, *20*(10), 977-989.

Huff, A. S. (1990). *Mapping strategic thought*. New York: John Wiley & Sons.

Ivakhnenko, G. G. (1971). Polynomial theory of complex systems. *IEEE Transactions on Systems, Man and Cybernetics*, *4*(3), 364-384.

Johnson-Laird, P. (1983). *Mental models*. Boston: Cambridge University Press.

Jonnassen, D. H., & Grabowski, B. (1993). *Handbook of individual differences, Learning and instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kearney, A. R., & Kaplan, S. (1997). Toward a methodology for the measurement of knowledge structures of ordinary people: The conceptual content cognitive map (3CM). *Environment and Behavior*, *29*(5), 579-617.

Kolodner, J. (1993). *Case-based reasoning*. San Mateo, CA: Morgan Kaufmann.

Levine, D., Stephan, D., Krehbiel, T., & Berenson, M. (2005). *Statistics for managers using Microsoft Excel, 4E*. Upper Saddle River, NJ: Prentice Hall.

Marakas, G. (2003). *Modern data warehousing, mining, and visualization: Core concepts*. Upper Saddle River, NJ: Prentice Hall.

Mohammed, S., Klimoski, R., & Rentsch, J. (2000). The measurement of team mental models: We have no shared schema. *Organizational Research Methods*, *3*(2), 123-165.

Müller, F., & Lemke, F. (2003). *Self-organizing data mining: Extracting knowledge from data.* Canada: Trafford Publishing.

Nonaka, I., & Konna, N. (1998). The concept of 'Ba': Building a foundation for knowledge creation. *California Management Review*, *40*(3), 40-54.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company*. New York: Oxford University Press.

Nunamaker, J. R., Briggs, R. O., Mittleman, D. D., Vogel, D. R., & Balthazard, P. A. (1997). Lessons from a dozen years of group support systems. *Journal of Management Information Systems, 13*, 333-360.

Powell, S. G., & Baker, K. R. (2007). *Management science: The art of modeling with spreadsheets* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Prince, G. M. (1970). *The practice of creativity*. NY: Harper & Row.

Reber, A. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*(3), 219-235.

Reisbeck, C., & Schank, R. (1989). *Inside case based reasoning*. Hillsdale, NJ: Erlbaum.

Robinson, S., Alifantis, T., Edwards, J. S., Ladbrook, J., & Waller, A. (2005). Knowledge-based improvement (KBI): Simulation and artificial intelligence for identifying and improving human decision-making in an operations system. *Journal of the Operational Research Society*, *54*, 912-921.

Sabherwal, R., & Becerra-Fernandez, I. (2003). An empirical study of the effect of knowledge management processes at individual, group, and organizational levels. *Decision Science*, *34*(2), 225-260.

Schoemaker, P. J. H. (1993). Multiple scenario development: Its conceptual and behavioral foundation. *Strategic Management Journal*, *14*, 193-213.

Schraagen, J. (1993). How experts solve a novel problem in experimental design. *Cognitive Science*, *17*, 285-309.

Seber, G., & Wild, C. (2003). *Nonlinear regression* (2nd ed.). New York: John Wiley & Sons.

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organizations*. New York: Doubleday/Currency.

Simon, H. A. (1974). How big is a chunk? *Science*, *183*, 482-488.

Simon, H. A. & Gobet, F. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Science*, *31*, 1-40.

Snee, R. D. (1973). Some aspects of nonorthogonal data analysis, part 1: Developing prediction equations. *Journal of Quality Technology*, *5*, 69-79.

Steiger, D. M. (1998). Enhancing user understanding in a decision support system: A theoretical basis and framework. *Journal of Management Information Systems, 15*(2), 199-220.

Steiger, D. M., & Sharda, R. (1996). Model analysis systems: Enhancing model analysis in decision support systems. *Information Systems Research, 7*(3), 328-341.

Steiger, N. M., & Steiger, D. M. (2007). Knowledge management in decision making. *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*. Computer Society Press, 10 pages.

Stephen, D. (2007). *PHStat2 v.2.7*. Upper Saddle River, NJ: Pearson Education.

Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge: A bottom-up model of skill learning. *Cognitive Science*, *25*(1), 203-230.

Sun, R., & Peterson, T. (1998). Some experiments with a hybrid model for learning sequential decision making. *Information Science*, *111*, 83-107.

Tsoukas, H. (2003). Do we really understand tacit knowledge? In M. Easterby-Smith & M. A. Lyles (Eds.), *The Blackwell handbook of organization learning and knowledge management* (pp. 410-427). Oxford, UK: Blackwell Publishing.

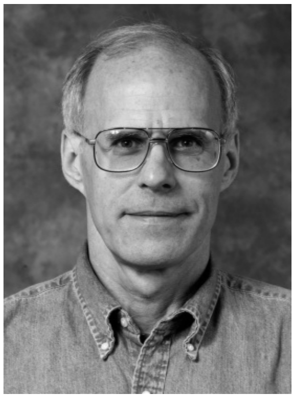
Ullman, L. (2001). *PHP and MySQL for dynamic web sites*. Berkley, CA: Peachpit Press.

Wagner, H. M. (1995). Global sensitivity analysis. *Operations Research*, *43*(6), 948-969.

Walsh, J. P. (1995). Managerial and organizational cognition: Notes from a trip down memory lane. *Organization Science*, *6*, 280-321.

Weick, K. E. (1990). Cartographic myths in organizations. In A. S. Huff (Ed.), *Mapping strategic thought* (pp. 1-10). New York: John Wiley and Sons.

Biographies



David M. Steiger is an Associate Professor of Management Information Systems at the University of Maine. He received his B.S. in Electrical Engineering and M.B.A. from the University of Texas at Austin and a Ph.D. in Management Science/Information Systems from Oklahoma State University. Between his M.B.A. and Ph.D. degrees, he spent 15 years in various analysis and managerial positions in industry, during which time he coauthored the winning paper in the 1986 INFORMS Edelman Award competition. Professor Steiger's research interests include decision support systems, knowledge management, model analysis systems, and inductive artificial intelligence technologies. His articles have appeared in *Information Systems Research, Management Science, Journal of Management Information Systems, Decision Support Systems, Interfaces, INFORMS Journal On Computing, European Journal of Operational Research, Knowledge Management Research & Practice,* and *Journal of Knowledge Management.*

Natalie M. Steiger, Ph.D. (North Carolina State, 1999) is an Associate Professor of Management at the University of Maine. Dr. Steiger was an engineering economic analyst and manager of petroleum-reserve information systems for Cities Service Oil and Gas Company and Occidental Petroleum Company from 1975 to 1986. Her research interests are the use of information and knowledge in decision-making; stochastic modeling; modeling of business/production systems, large-scale simulation modeling and analysis, including development of techniques to improve simulation analysis in practical applications; and other applications of operations research. She has published in the *INFORMS Journal On Computing*, *IIE Transactions*, *Management Science*, *ACM TOMACS,* and *Journal of Simulation.* She is a member of IIE, INFORMS, and the INFORMS Simulation Society.