UNIVERSITI SAINS MALAYSIA

Stamford College

Second Semester Examination
Academic Session 2007/2008
April 2008

External Degree Programme
Bachelor of Computer Science (Hons.)

CMT314 – Decision Support Systems & Intelligent Systems

Duration : 2 hours

INSTRUCTIONS TO CANDIDATE:

- Please ensure that this examination paper contains **THREE** questions in **SIX** printed pages before you begin the examination.
- Answer **ALL** questions.
- On each page, write *only your **Index Number***.

1. (a) (i) Briefly explain the major activities in the implementation phase of a decision-making process.

(10/100)

(ii) Your company is considering opening a branch in Indonesia. As the ICT Business Development Manager, list the typical activities in each phase of the decision-making process whether to open or not to open the branch.

(20/100)

(b) (i) Describe the major characteristics of a decision support system (DSS).

(10/100)

(ii) Explain how personality type, cognitive style and decision style might affect the process of decision-making.

(20/100)

(c) (i) How does Online Analytical Processing (OLAP) differ from Online Transaction Processing (OLTP)?

(10/100)

(ii) Describe 'multidimensional analysis' and explain its potential benefits to the knowledge workers.

(30/100)

2. (a) (i) Briefly explain the following terms:

- Knowledge
- Business Intelligence
- Business Analytics
- Collaborative Computing

(10/100)

(ii) Based on your understanding of the Case Application on 'Clay Processing Planning' as in Appendix, discuss the possible use of a similar decision support system by organizations processing other non-clay materials.

(20/100)

(iii) Briefly describe **two (2)** new research directions in decision support systems.

(20/100)

(b) Explain the differences between the following:

(i) Protocol analysis and observation

(ii) Forward chaining and backward chaining

(iii) Supervised learning and unsupervised learning

(30/100)

(c) Consider the following passage:

"When it is a Tuesday and if it is 10am, I have a class to attend. I bring along an umbrella to class when there are dark clouds in the sky and if I have an umbrella. If I do not have an umbrella, I would borrow one. After borrowing one, I would then have an umbrella to bring along to the class. There are dark clouds during the rainy season."

Present the knowledge found in the passage using a suitable knowledge representation format.

(20/100)

3. (a) Describe the motivations for having an explanation subsystem in an expert system. What are the different kinds of explanations that can be generated? How are explanations generated?

(30/100)

(b) How does a neural network learn? In your answer, assume that there is a single processing element and that supervised learning is employed. Draw a diagram to illustrate your answer.

(20/100)

(c) What is knowledge discovery? Briefly describe the knowledge discovery process.

(24/100)

(d) Can computers think? Are computers intelligent? Discuss your answer.

(26/100)

Appendix

CLAY PROCESS PLANNING AT IMERYS: A CLASSICAL CASE OF DECISION MAKING

Part 3: The Process Optimization Model

INTRODUCTION

This case application continues the effort described in Case Applications 2.1 and 2.2. The Process OPtimization (POP) development team at English China Clay International (ECCI, which became IMERYS) in Sandersville, Georgia, developed a large-scale mathematical programming model that describes its clay processing operation from the mines to the finished product. Here we describe the structure of the POP model: a large-scale, generalized, multicommodity network flow model with side constraints. We further describe how the data and model are managed. Finally, we describe how the model is and will be used. The prototyping development process followed in developing the POP DSS is described in detail in Case Application 6.1.

THE PLANTS

The scope of the first phase of the project originally called for developing an integrated model representing four plants—two hydrous plants, a large calcine plant, and a small calcine plant—but did not represent the mines (calcine is dry clay, and hydrous has more moisture; different products are made from each, and almost any set of clays can be blended to generate a final product with unique properties). The mining portion of the model was added later. While development of the model for the small calcine plant was underway, ECCI was purchased by IMATEL (France), and eventually one hydrous plant, about one-fifth of the large calcine plant, and the small calcine plant were sold per a U.S. Justice Department ruling. As outlined in Case Application 6.1, we had completed development of the model of both calcine plants at that time. For validation purposes, we kept the plants in the model until it became operational. The POP DSS model deployed in late 1999 represented one hydrous plant, the large calcine plant, and the small calcine plant. Later, we replaced the small plant with external market purchases and demands for intermediate clays that were shipped to it.

THE MODEL BUILDING BLOCKS

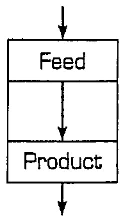
The decision variables include which mines to excavate, how much and what kind of crude clays to extract from each one, how to blend crude clays, which equipment to process the clay on, the speed at which to process the clay, what intermediate blends (recipes) to use, what final blends to use, what demands to meet (or not meet if necessary), what final clays to purchase from the open market, and so on.

Fortunately, the multicommodity network flow problem represents flow problems of many commodities (e.g., different clays) through common links (arcs) that generally have capacity limits. The model can be represented graphically, making it easy to sketch and understand. Ours is a generalized model; that is, each link that allows flow has a multiplier (a recovery factor for a process) between 0 and 1 indicating how much of the flow actually reaches the node at the end of the link. This is used to model *losses* that result from chemical and physical transformation of the clays. In addition, there are some side constraints that enforce blends and enforce mutual capacities on the links (e.g., the total flow through each arc for all commodities cannot exceed the capacity in terms of flow or time). This is a static model.

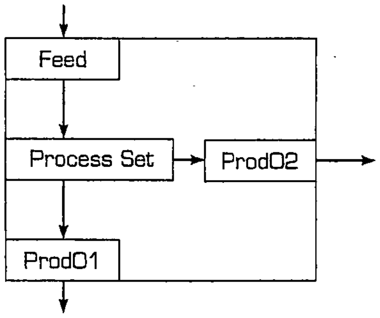
Developing a standard set of building blocks made it easier for the team to develop and implement the model. Given a particular clay, there are several model building blocks, but the most important one is the process. These are entities that represent a type of equipment processing the clay. For example, transporting the clay from a mine to a particular plant is a process. Another process is grinding. Other building blocks, such as a holding tank, follow naturally from the process definition. Some processes are simply represented as a pair of nodes: a source (a supply, e.g., a mine), a sink (the demand for a finished clay), and a link that allows flow between pairs of building blocks. Every process has a set of clays that can flow through it. For each clay flowing through a process, the following data must be specified: the rate of flow (in tons per hour, which varies by clay), a unit cost per ton for processing, a unit cost per hour utilized, a recovery factor (the multiplier between 0 and 1), a capacity limit on the flow, and a capacity limit on the processing time.

The basic building block of a simple process consists of two nodes and a single arc. The first node is the feed node. Any preceding processes can feed the clay into the process through this node. The second node is the product node. This is where the processed clay arrives and is ready for transport to its next destination. The decision variable is to determine the flow through the process (on the arc). A simple process looks like

Appendix



Complex processes have two or more distinct products (e.g., a categorizing process divides clay into small and large particle sizes, each of which is processed differently afterward; so each product has a different recovery factor, while the rate and unit costs of processing are unchanged. A complex process has an intermediate node (the process set node), a product node for each one, and arcs to link them. It looks like



Chemicals that alter the clays' properties are added to the clays in different processes. The amount used is proportional to the flow (in pounds per ton), and, depending on the rate the process uses, different chemical amounts can be involved. Alternative processing for the same clay may lead to the use of different chemicals.

Clays can flow from plant to plant, from the *economy* into the plant, from mine to plant, and so on. The model is then built by connecting these processes with arcs that transport the clays. These arcs represent any transporting of clays. There are about 15 crude clays, five families of hydrous clays, and three main calcine products. Though few in number, these clays can be combined with each other and with clays obtained from other plants or on the open market to produce several hundred different final clays. There are hundreds of ways to blend the crudes to form any of the hydrous family clays. Each family goes through the production process in several different ways. There are different ways to process each particular clay, and different blends and chemical amounts can be used. The model was to determine the optimal blends to use.

The model, when solved, determines the clay flows (decision variables, in tons) and the time consumed for each clay in each process. These values are capacitated, and the total flow and total time consumed are also capacitated, both because of physical limitations of the processing equipment and the required characteristics of the finished products. The recipes used and which processes are running at capacity are of great interest to the company for planning purposes. The mining operations are also a "process," as is meeting the demand for each clay.

The objective is to maximize profit. Each finished product clay has a unit price for every form of it that is sold (slurry, bulk, bag, etc.). More than 2.3 million tons of crude clay processing annually was modeled.

MODELING DIFFICULTIES

What made this model difficult to construct and interesting was the large size (initially more than 8,000 constraints and 35,000 variables; by 2003 there were over 80,000 constraints and 170,000 variables) and the fact that several different process characteristics were estimated because the processes had not yet been constructed. There were also points in the processing where by-products were fed back into the system to an earlier step (clay recoveries).

Once the small calcine plant and a portion of the large calcine plant were sold, the flows into these portions of the model were turned off by setting the capacity of the calcining process equal to zero, and open market purchases were added for some final products. A second hydrous plant was never modeled. Later, the size of the model increased by 50 percent as other plants and clays were added.

THE LINGO MODELING LANGUAGE AND THE ACCESS DATABASE INTEGRATION

The model was developed in Lingo (a modeling language from Lindo Systems Inc., lindo.com), which integrates directly with a Microsoft Access database of more than 10 relational tables through the Microsoft @ODBC interface. The Lingo model lines are specified independently from the data link statements (links). The Process OPtimization Lingo model is populated with data from the database, generates the model, solves it, and loads the solution directly back into the database automatically. Lingo model lines generally look like shorthand for the algebra of mathematical programming, thus providing a familiar vehicle for model building. For example, the Lingo model line for the supply constraints of a transportation problem (from factories to customers) might be

```
@FOR(FACTORY( I):
        @SUM(CUSTOMER( J): FLOW( I, J))
            <= CAPACITY( I));
```

which means: For every FACTORY(I), SUM all the flows from supply node I to demand node J over all CUSTOMER(J) (all customers), (FLOW(I,J)), and set that value to be less than or equal to the available CAPACITY(I) at FACTORY(I). There are special data statements

Appendix

specifying all necessary data to identify the sets FACTORY, CUSTOMER, and CAPACITY. The POP model's mining portion looks very much like a modified transportation problem. Limits on blends can be specified (e.g., clay B must constitute between 80 and 95 percent of the blend).

POP DSS USE

The DSS, written as a menu-oriented Access database table, manages the data in the system. A particular scenario is set up in the Access tables through a friendly graphical user interface (GUI) screen. The user sets the demands, makes other adjustments to the processes, and then activates Lingo with the click of a button. Lingo automatically generates the entire model from its compact representation and the data as specified in the database, and solves it. Lingo loads its solution back into the database and returns control to the menu-oriented GUI. Access programs then produce managerially meaningful graphs of utilization and reports on clay extraction and processing. Trouble spots are identified, the case can be saved, and another scenario can be run.

For a fixed time period (one year, one quarter, two weeks, etc.), the solution to the model indicates which mines are active, how much clay is mined from each mine, to which processing unit the clay is shipped from the mines, and the appropriate crude blends (recipes) to be used. It determines all the clay flows throughout the entire system and which clays to purchase from the market. The model quickly identifies which processes are running at capacity and indicates the potential increase in profit that could be obtained if these capacities could be increased (through sensitivity analysis). Sometimes there are underutilized processes that could handle some of the load of the limited processes but are somewhat inefficient at doing so. Plant managers are reluctant to use these processes but carefully examine them and sometimes activate them.

The model also indicates how to handle the situation now that some of the higher-quality clay mines are depleted and new processes have been introduced. Finally, underutilization of some processes indicates that some final products, normally produced at other plants (not yet in the model), could be produced at the plants represented by the model. Several of these clays have already been added to POP.

The most interesting aspect of the model is that the engineers and managers who structure the plants were doing an excellent job of keeping them fine-tuned without access to these analytical tools. The model did recommend using different mines from time to time, and it has provided guidance on how to manage the mines for ten years. The total amount of clay being processed is about the same as what the model solution recommends, which certainly helped to validate the model. What the model is best used for is determining how to handle the resources that are 100 percent utilized (bottlenecks) and how to handle new and unexpected situations, such as new clays, new demands, and new processes. It also provides answers quickly and easily, thus guiding managers and engineers in their decision-making. When a plant was closed in 2001, the production and demand for its products were moved to the main hydrous plant that was already in the POP DSS. POP accurately determined the blends that indicated how to handle this record throughput optimally. Even the plant manager, initially skeptical, agreed that his plant could handle the load once he saw POP's recommendations.

As mentioned in Case Application 2.2, the cost of operating a new process was determined, thus guiding budgeting decisions for the next fiscal year. The model is used for annual planning. It is also used in the short term for scheduling specific large orders in with the forecasted demands. Essentially, the POP DSS is used for strategic planning (1–5 years), tactical planning (3–6 months), and operational planning (2 weeks). A simple factor is changed to generate a model that spans any needed time frame.

- oooOooo -