

4

The Model Formulation Process

*Count what is countable, measure what is measurable, and
what is not measurable, make measurable.
Galileo Galilei(1564-1642)*

4.1 The Overall Process

In using any kind of analytical or modeling approach for attacking a problem, there are five major steps:

- 1) Understanding the real problem.
- 2) Formulating a model of the problem.
- 3) Gathering and generating the input data for the model (e.g., per unit costs to be used, etc.).
- 4) Solving or running the model.
- 5) Implementing and interpreting the solution in the real world.

In general, there is a certain amount of iteration over the five (e.g., one does not develop the most appropriate model the first time around). Of the above, the easiest is the solving of the model on the computer. This is not because it is intrinsically easiest, but because it is the most susceptible to mathematical analysis. Steps 1, 3, and 5 are, if not the most difficult, at least the most time consuming. Success with these steps depends to a large extent upon being very familiar with the organization involved (e.g., knowing who knows what the real production rate is on the punch press machine). Step 2 requires the most analytical skill. Steps 1 and 5 require the most people skills.

Formulating good models is an art bordering on a science. The artistic ability is in developing simple models that are nevertheless good approximations of reality. We shall see that there are a number of classes of problems that are well approximated by optimization models.

With all of the above comments in mind, we will devote most of the discussion to formulation of optimization models, stating what universal truths seem to apply for steps (3) and (5), and giving an introduction to the mechanics of step (4).

4.2 Approaches to Model Formulation

We take two approaches to formulating models:

- 1) Template approach,
- 2) Constructive approach.

The constructive approach is the more fundamental and general. However, readers with less analytic skill may prefer the template approach. The latter is essentially a “model in a can” approach. In this approach, examples of standard applications are illustrated in substantial detail. If you have a problem that closely resembles one of these “template” models, you may be able to adjust it to your situation by making modest changes to the template model. The advantage of this approach is that the user may not need much technical background if there is a template model that closely fits the real situation.

4.3 The Template Approach

You may feel more comfortable and confident in your ability to structure problems if you have a classification of “template” problems to which you can relate new problems you encounter. We will present a classification of about a half dozen different categories of problems. In practice, a large real problem you encounter will not fit a single template model exactly, but might require a combination of two or more of the categories. The classification is not exhaustive, so you may encounter or develop models that seem to fit none of these templates.

4.3.1 Product Mix Problems

Product mix problems are the problem types typically encountered in introductory LP texts. There are a collection of products that can be sold and a finite set of resources from which these products are made. Associated with each product are a profit contribution rate and a set of resource usage rates. The objective is to find a mix of products (amount of each product) that maximizes profit, subject to not using more resources than are available.

These problems are always of the form “Maximize profit subject to less-than-or-equal-to constraints”.

4.3.2 Covering, Staffing, and Cutting Stock Problems

Covering, staffing, and cutting stock problems are complementary (in the jargon, they are called dual) to product mix problems in that their form is “Minimize cost subject to greater-than-or-equal-to constraints”. The variables in this case might correspond to the number of people hired for various shifts during the day. The constraints arise from the fact that the mix of variables chosen must “cover” the requirements during each hour of the day.

4.3.3 Blending Problems

Blending problems arise in the food, feed, metals, and oil refining industries. The problem is to mix or blend a collection of raw materials (e.g., different types of meats, cereal grains, or crude oils) into a finished product (e.g., sausage, dog food, or gasoline). The cost per unit of the finished product is minimized and it is subject to satisfying certain quality constraints (e.g., percent protein \geq 15 percent).

4.3.4 Multiperiod Planning Problems

Multiperiod planning problems constitute perhaps the most important class of models. These models take into account the fact that the decisions made in this period partially determine which decisions are allowable in future periods. The submodel used each period may be a product mix problem, a blending problem, or some other type. These submodels are usually tied together by means of inventory variables (e.g., the inventory of raw materials, finished goods, cash, or loans outstanding) that are carried from one period to the next.

4.3.5 Network, Distribution, and PERT/CPM Models

Network LP models warrant special attention for two reasons: (a) they have a particularly simple form, which makes them easy to describe as a graph or network, and (b) specialized and efficient solution procedures exist for solving them. They, therefore, tend to be easier to explain and comprehend. Network LPs frequently arise from problems of product distribution. Any enterprise producing a product at several locations and distributing it to many customers may find a network LP relevant. Large problems of this type may be solved rapidly by the specialized procedures.

One of the simplest network problems is finding the shortest route from one point in a network to another. A slight variation on this problem, finding the longest route, happens to be an important component of the project management tools PERT (Program Evaluation and Review Technique) and CPM (Critical Path Method).

Close cousins of network models are input/output and vertically integrated models. General Motors, for example, makes engines in certain plants. These engines might be sold directly to customers, such as industrial equipment manufacturers, or the engines may be used in GM's own cars and trucks. Such a company is said to be vertically integrated. In a vertically integrated model, there is usually one constraint for each type of intermediate product. The constraint mathematically enforces the basic law of physics that the amount used of an intermediate product by various processes cannot exceed the amount of this product produced by other processes. There is usually one decision variable for each type of process available.

If one expands one's perspective to the entire economy, then the models considered tend to be similar to the input/output model popularized by Wassily Leontief (1951). Each industry is described by the input products required and the output products produced. These outputs may in turn be inputs to other industries. The problem is to determine appropriate levels at which each industry should be operated in order to satisfy specific consumption requirements.

4.3.6 Multiperiod Planning Problems with Random Elements

One of the fundamental assumptions of optimization models is that all input data are known with certainty. There are situations, however, where certain key data are highly random. For example, when an oil company makes its fuel oil production decisions for the coming winter, the demand for that fuel oil is very much a random variable. If, however, the distribution probabilities for all the random variables are known, then there is a modeling technique for converting a problem that is an optimization model, except for the random elements, into an equivalent, although possibly larger, deterministic optimization model. Such models are sometimes called stochastic programs.

4.3.7 Financial Portfolio Models

An important application of optimization in the last ten years has been in the design of financial investment portfolios. In its simplest form, it is concerned with how much to invest in a collection of risky investments, so that a good compromise is struck between a high expected return and a low risk.

More complicated applications of this idea are concerned with investing so as to track some popular financial index, such as the S&P 500.

4.3.8 Game Theory Models

Game theory is concerned with the analysis of competitive situations. In its simplest form, a game consists of two players, each of whom has available to them a set of possible decisions. Each player must choose a strategy for making a decision in ignorance of the other player's choice. Some time after a decision is made, each player receives a payoff that depends on which combination of decisions was made. The problem of determining each player's optimal strategy can be formulated as a linear program.

Not all problems you encounter will fit into one of the above categories. Many problems will be combinations of the above types. For example, in a multiperiod planning problem, the single period subproblems may be product mix or blending problems.

4.4 Constructive Approach to Model Formulation

The constructive approach is a set of guidelines for constructing a model from the ground up. This approach requires somewhat more analytical skill, but the rules apply to any situation you are trying to model. The odds are low you will find a template model that exactly matches your real situation. In practice, a combination of these two approaches is needed.

For the constructive approach, we suggest the following three-step approach for constructing a model, which, with apologies to Sam Savage, might be called the ABC's of modeling:

- A. Identify and define the decision variables or *Adjustable cells*. Defining a decision variable includes specifying the units in which it is measured (e.g., tons, hours, etc.). One way of trying to deduce the decision variables is to ask the question: What should be the format of a report that gives a solution to this problem? (For example, the numbers that constitute an answer are: the amount to produce of each product and the amount to use of each ingredient.) The cells in this report are the decision variables.
- B. Define how we measure *Best*. More officially, define our objective or criterion function, including the units in which it is measured. Among useable or feasible solutions, how would preference/goodness (e.g., profit) be measured?
- C. Specify the *Constraints*, including the units in which each is measured. A way to think about constraints is as follows: Given a purported solution to a problem, what numeric checks would you perform to check the validity of the solution?

The majority of the constraints in most problems can be thought of as *sources-equals-uses* constraints. Another common kind of constraint is the *definitional* or *accounting* constraint. Sometimes the distinction between the two is arbitrary. Consider a production setting where we: *i*) start with some beginning inventory of some commodity, *ii*) produce some of that commodity, *iii*) sell some of the commodity, and *iv*) leave some of the commodity in ending inventory. From the sources-equals-uses perspective, we might write:

$$\text{beginning inventory} + \text{production} = \text{sales} + \text{ending inventory}.$$

From the definitional perspective, if we were thinking of how ending inventory is defined, we would write:

$$\text{ending inventory} = (\text{beginning inventory} + \text{production}) - \text{sales}.$$

The two perspectives are in fact mathematically equivalent.

For any application, it is useful to do each of the above in words first. In order to illustrate these ideas, consider the situation in the following example.

4.4.1 Example

Deglo Toys has been manufacturing a line of precision building blocks for children for a number of years. Deglo is faced with a standard end-of-the-year problem known as the build-out problem. It is about to introduce a new line of glow-in-the-dark building blocks. Thus, they would like to deplete their old-technology inventories before introducing the new line. The old inventories consist of 19,900 4-dimple blocks and 29,700 8-dimple blocks. These inventories can be sold off in the form of two different kits: the Master Builder and the Empire Builder. The objective is to maximize the revenue from the sale of these two kits. The Master kit sells for \$16.95, and the Empire kit sells for \$24.95. The Master kit is composed of 30 4-dimple blocks plus 40 8-dimple blocks. The Empire kit is composed of 40 4-dimple blocks plus 85 8-dimple blocks. What is an appropriate model of this problem?

4.4.2 Formulating Our Example Problem

The process for our example problem would be as follows:

- a) The essential decision variables are:
 M = number of master builder kits to assemble and
 E = number of empire builder kits to assemble.
- b) The objective function is to maximize sales revenue (i.e., Maximize $16.95M + 24.95E$).
- c) If someone gave us a proposed solution (i.e., values for M and E), we would check its feasibility by checking that:
 - i. the number of 4-dimple blocks used $\leq 19,900$ and
 - ii. the number of 8-dimple blocks used $\leq 29,700$.

Symbolically, or algebraically, this is:

$$\begin{aligned} 30M + 40E &\leq 19,900 \\ 40M + 85E &\leq 29,700 \end{aligned}$$

In LINGO form, the formulation is:

```
MAX = 16.95 * M + 24.95 * E;
      30 * M   + 40 * E <= 19900;
      40 * M   + 85 * E <= 29700;
```

with solution:

```
Optimal solution found at step:          0
Objective value:                      11478.50
Variable          Value          Reduced Cost
M                  530.0000          0.0000000
E                  100.0000          0.0000000
Row   Slack or Surplus   Dual Price
  1         11478.50         1.000000
  2          0.000000         0.4660526
  3          0.000000         0.7421052E-01
```

Thus, we should produce 530 Master Builders and 100 Empire Builders.

4.5 Choosing Costs Correctly

Choosing costs and profit contribution coefficients in the objective requires some care. In many firms, cost data may not be available at the detailed level required in an optimization model. If available, the “official” cost coefficients may be inappropriate for the application at hand.

The basic rule is fairly simple: The cost coefficient of a variable should be the rate of change of the total cost as the variable changes. We will discuss the various temptations to violate this rule. The two major temptations are sunk costs and joint costs.

4.5.1 Sunk vs. Variable Costs

A *sunk cost* is a cost that has already been incurred or committed to, although not necessarily paid. A *variable cost* is a cost that varies with some activity level. Sunk costs should not appear in any coefficient of a decision variable. Whether a cost is sunk or variable depends closely upon the length of our planning horizon. A general rule is that: In the short run, all costs are sunk, while all costs are variable in the long run. The following example illustrates.

Sunk and Variable Cost Example

A firm prepared a profit contribution table for two of its products, X and Y :

<u>Product:</u>	<u>X</u>	<u>Y</u>
Selling price/unit	\$1000	\$1000
Material cost/unit	\$200	\$300
Labor cost/unit	<u>\$495</u>	<u>\$300</u>
Net Profit contribution	\$305	\$400

These two products use a common assembly facility that has a daily capacity of 80 units. Product specific production facilities limit the daily production of X to 40 units and Y to 60 units. The hourly wage in the company is \$15/ hour for all labor. The obvious model is:

$$\begin{aligned} \text{Max} &= 305 * X + 400 * Y; \\ X &\leq 40; \\ Y &\leq 60; \\ X + Y &\leq 80; \end{aligned}$$

The solution is to produce 60 Y 's and 20 X 's. At \$15 per hour, the total labor required by this solution is $20 \times 495/15 + 60 \times 300/15 = 1860$ hours per day.

Now, let us consider some possible additional details or variations of the above situation. Some firms, such as some automobile manufacturers, have had labor contracts that effectively guarantee a job to a fixed number of employees during the term of the contract (e.g., one year). If the above model is being used to decide how many employees to hire and commit to before signing the contract, then the \$15/hour used above is perhaps appropriate, although it may be too low. In the U.S., the employer also must pay Social Security and Medicare taxes that add close to 8% to the labor bill. In addition, the employer typically also covers the cost of supplemental health insurance for the employee, so the cost of labor is probably closer to \$20 per hour rather than \$15.

Once the contract is signed, however, the labor costs then become sunk, but we now have a constraint that we can use at most 1860 hours of labor per day. The variable profit contributions are now:

Product:	<u>X</u>	<u>Y</u>
Selling price/unit	\$1000	\$1000
Material cost/unit	<u>\$200</u>	<u>\$300</u>
Net Profit contribution	\$800	\$700

Now, X is the more profitable product. Before we jump to the conclusion that we should now produce 40 X 's and 40 Y 's, we must recall that labor capacity is now fixed. The proper, short term, model is now:

$$\begin{aligned} \text{MAX} &= 800 * X + 700 * Y; \\ X &\leq 40; \\ Y &\leq 60; \\ X + Y &\leq 80; \\ 33 * X + 20 * Y &\leq 1860; \end{aligned}$$

with solution:

Optimal solution found at step:		1
Objective value:		58000.00
Variable	Value	Reduced Cost
X	20.00000	0.000000
Y	60.00000	0.000000
Row	Slack or Surplus	Dual Price
1	58000.00	1.000000
2	20.00000	0.000000
3	0.000000	215.1515
4	0.000000	0.000000
5	0.000000	24.24242

Therefore, we still produce the same mix of products.

Now, suppose in order to be competitive, the selling price of X must be dropped by \$350 to \$650. Also, we still have our labor contract that says we may use up to and must pay for all of 1860 hours of labor per day. The correct model is:

$$\begin{aligned} \text{Max} &= 450 * X + 700 * Y; \\ X &\leq 40; \\ Y &\leq 60; \\ X + Y &\leq 80; \\ 33 * X + 20 * Y &\leq 1860; \end{aligned}$$

with still the same solution of $X = 20$ and $Y = 60$. If we (incorrectly) charge for labor, however, the model is:

$$\begin{aligned} \text{Max} &= -45 * X + 400 * Y; \\ X &\leq 40; \\ Y &\leq 60; \\ X + Y &\leq 80; \\ 33 * X + 20 * Y &\leq 1860; \end{aligned}$$

and we would incorrectly conclude that X should not be produced.

There are many planning situations similar to the above. For example, an airline or a trucking firm may use essentially the same model for long-range fleet sizing decisions as for daily fleet routing decision. When solving the long-term fleet sizing decision, the cost of capital should be included in the daily cost of having a vehicle. On the other hand, when making short-run routing decisions, the cost of capital should not be included in the daily cost of a vehicle. However, the number of vehicles used is constrained to be no greater than the fleet size chosen in the long-term plan. Only operating costs that vary with the amount of usage of the vehicle should be included when solving the short-term model.

4.5.2 Joint Products

We say we have joint products or byproducts if a single process produces several products. The key feature is that, if you run the process in question, you unavoidably get some amount of each of the joint products. Some examples are:

Process	Joint Products
Crude oil distillation	gasoline, oil, kerosene, tar
Raw milk processing	whole milk, skim milk, 2%, cream, yogurt
Meat processing	light meat, dark meat, steak, chuck roast
Semi-conductor manufacturing	high speed chips, low speed chips
Mining of precious metal ore	gold, silver, copper
Sales calls	sales of various products in product line

There is a temptation, perhaps even a requirement by taxing authorities, that the cost of the joint process be fully allocated to the output products. The important point is that, for decision-making purposes, this allocation serves no purpose. It should be avoided. The proper way to model a joint product process is to have a separate decision variable for each output product, and a decision variable for the joint production process. Costs and revenues should be applied to their associated decision variables (e.g., the cost of distillation should be associated with the decision variable of how much crude to distill). The fact that, if you want to produce gasoline, then you must incur the cost of distillation is taken care of by the constraints. Let us illustrate with an example.

Joint Cost Example

The Chartreuse Company (CC) raises pumpkins. It costs \$800 to plant, harvest and sort a ton of raw pumpkins. CC has capacity to plant and harvest 150 tons of pumpkins. In spite of CC's best efforts at genetic engineering, harvested pumpkins fall equally into three classes of pumpkins of increasing quality: Good, Premium, and Exquisite. Once sorted, it costs \$100 per ton to get each of the classes ready for market. Alternatively, pumpkins from any class can be discarded at zero additional cost. Prices have dropped recently, so there is concern about whether it is profitable to sell all grades of pumpkins. Current selling prices per ton for the three grades are: \$700, \$1100, and \$2200. How much should be processed and sold of each grade?

A proper model is:

$$\begin{aligned} \text{MAX} &= (700 - 100) * G + (1100 - 100) * P + (2200 - \\ &100) * E - 800 * R; \\ R &\leq 150; \\ G &\leq .3333333 * R; \\ P &\leq .3333333 * R; \\ E &\leq .3333333 * R; \end{aligned}$$

With solution:

Optimal solution found at step:	0	
Objective value:	65000.0	
Variable	Value	Reduced Cost
G	50.00000	0.0000000
P	50.00000	0.0000000
E	50.00000	0.0000000
R	150.0000	0.0000000
Row	Slack or Surplus	Dual Price
1	65000.00	1.000000
2	0.0000000	433.3333
3	0.0000000	600.0000
4	0.0000000	1000.000
5	0.0000000	2100.000

There is a temptation to allocate the cost of planting, harvesting and sorting, over all three grades to get the model:

$$\begin{aligned} \text{MAX} &= (700 - 100 - 2400/3) * G + (1100 - 100 - 2400/3) * P + (2200 - 100 - 2400/3) * E ; \\ G &\leq .3333333 * 150; \\ P &\leq .3333333 * 150; \\ E &\leq .3333333 * 150; \end{aligned}$$

Given their (apparent) negative profit contribution in the above model, good pumpkins will not be produced. If we then allocate the planting, harvesting, and sorting costs over just P and E , we get:

$$\begin{aligned} \text{MAX} &= (1100 - 100 - 2400/2) * P + (2200 - 100 - 2400/2) * E; \\ G &\leq .3333333 * 150; \\ P &\leq .3333333 * 150; \\ E &\leq .3333333 * 150; \end{aligned}$$

Now, of course, Premium grade is not worth producing. This leaves the Exquisite grade to carry the full cost of planting, harvesting, and sorting, and then we see it is not worth producing. Thus, even though we started with a profitable enterprise, blind use of allocation of joint costs caused us to quit the profitable business. The moral to the story is to not do cost allocation.

4.5.3 Book Value vs. Market Value

A common problem in formulating an optimization model for decisionmaking is what cost should be attached to product that is used from inventory. A typical accounting system will carry a book value for product in inventory. The temptation is to use this readily available number as the cost of using product from inventory. For example, suppose you are a gasoline distributor who bought 10,000 gallons of Regular gasoline last month for \$2.77 per gallon. Due to unforeseen events, this month you still have 5,000 gallons of that Regular gasoline in inventory. Now the market price for Regular gasoline has dropped to \$2.70 per gallon, and you are contemplating your production and market operations for this month. How much should you charge yourself for the use of this Regular in inventory? One person might argue that the purchase is now a sunk cost so we should charge ourselves 0. Others might argue that proper "Accounting" says we should charge the book value, \$2.77/gallon. Which is it? The simple quick answer is that for decision making purposes, book value should always be disregarded, except when required by law for the calculation of taxes. Material in

inventory should be treated as having zero cost, however, you should completely enumerate all possible options of what you can do with this inventory, including selling it on the open market.

It helps to clarify issues by completing all the details for our little example and explicitly defining decision variables for all the possible actions available. You can buy or sell Regular in unlimited amounts this month for \$2.70/gallon, however, it costs you \$0.01/gallon in transportation and transaction costs for any gasoline you buy to get it onto your property. Similarly, for any gasoline you sell, there is a transaction cost of \$0.02 per gallon. What can be done with Regular gasoline? It can be sold directly, or it can be blended in equal proportions with Premium gasoline to produce Midgrade gasoline. You have one customer who is willing to pay \$2.82/gallon of Midgrade delivered to his door for up to 6000 gallons, and a second customer who is willing to pay \$2.80/gallon of Midgrade delivered to his door for up to 8000 gallons. Premium gasoline can be purchased in unlimited amounts for \$2.90/gallon. What should we do with our Regular gasoline: nothing, sell it back to the market, buy some Premium to blend with Regular (perhaps even buying more Regular) to sell to customer 1, to customer 2? Following the ABC's of optimization, step A is to define our decision variables: RB = gallons of additional Regular gasoline bought on the market this month, RS = gallons of Regular directly sold on the market this month, PB = gallons of Premium bought, $MS1$ = gallons of Midgrade sold to customer 1, and $MS2$ = gallons of Midgrade sold to customer 2. Step B, the objective function is to maximize revenues minus costs. Step C is to specify the constraints. The two main constraints are the "Sources Equal Uses" constraints for Regular and Premium. A formulation is given below. Recall that a gallon of Midgrade uses a half gallon of Regular and half gallon of Premium. To make the solution report easier to understand, we have given a [row name] to each constraint.

```
!Maximize revenues - costs;
MAX = (2.70 - .02)*RS + (2.82 - .02)*MS1 + (2.80-.02)*MS2
      - (2.70 + .01)*RB - (2.90 + .01)*PB;
!Sources = uses for Regular and Premium;
[SEQUR] 5000 + RB = RS + .5*(MS1 + MS2);
[SEQUP] PB = .5*(MS1 + MS2);
!Upper limits on amount we can sell;
[UL1] MS1 <= 6000;
[UL2] MS2 <= 8000;
```

Notice there is no explicit charge for Regular in inventory. The book value of \$2.77 appears nowhere in the formulation. Inventory is treated as a sunk cost or free good, however, we have included the option to sell it directly. Thus, using Regular to blend Midgrade must compete with simply selling the Regular directly at the current market price. A solution is:

Objective value: 13430.00		
Variable	Value	Reduced Cost
RS	2000.000	0.000000
MS1	6000.000	0.000000
MS2	0.000	0.015000
RB	0.000	0.030000
PB	3000.000	0.000000
Row	Slack or Surplus	Dual Price
1	13430.000	1.000000
SEQUR	0.000	-2.680000
SEQUP	0.000	-2.910000
UL1	0.000	0.005000

UL2

8000.000

0.000000

Thus, it is more profitable to blend the Regular inventory with Premium to sell it to customer 1 than to sell it directly to the market, however, selling Regular directly back to the market is more profitable than selling to customer 2 blended into Midgrade.

4.6 Common Errors in Formulating Models

When you develop a first formulation of some real problem, the formulation may contain errors or bugs. These errors will fall into the following categories:

- A. Simple typographical errors;
- B. Fundamental errors of formulation;
- C. Errors of approximation.

The first two categories of errors are easy to correct once they are identified. In principle, category *A* errors are easy to identify because they are clerical in nature. In a large model, however, tracking them down may be a difficult search problem. Category *B* errors are more fundamental because they involve a misunderstanding of either the real problem or the nature of LP models. Category *C* errors are subtler. Generally, a model of a real situation involves some approximation (e.g., many products are aggregated together into a single macro-product, the days of a week are lumped together, or costs that are not quite proportional to volume are nevertheless treated as linear). Avoiding category *C* errors requires skill in identifying which approximations can be tolerated.

With regard to category *A* errors, if the user is fortunate, category *A* errors will manifest themselves by causing solutions that are obviously incorrect.

Errors of formulation are more difficult to discuss because they are of many forms. Doing what we call dimensional analysis can frequently expose the kinds of errors made by a novice. Anyone who has taken a physics or chemistry course would know it as “checking your units.” Let us illustrate by considering an example.

A distributor of toys is analyzing his strategy for assembling Tinkertoy sets for the upcoming holiday season. He assembles two kinds of sets. The “Big” set is composed of 60 sticks and 30 connectors, while the “Tot” set is composed of 30 sticks and 20 connectors. An important factor is, for this season, he has a supply of only 60,000 connectors and 93,000 sticks. He will be able to sell all that he assembles of either set. The profit contributions are \$5.5 and \$3.5 per set, respectively, for Big and Tot. How much should he sell of each set to maximize profit?

The distributor developed the following formulation. Define:

- B = number of Big sets to assemble;
- T = number of Tot sets to assemble;
- S = number of sticks actually used;
- C = number of connectors actually used.

$$\begin{aligned} \text{MAX} &= 5.5 * B + 3.5 * T; \\ B - 30 * C - 60 * S &= 0; \\ T - 20 * C - 30 * S &= 0; \\ &C \leq 60000; \\ &S \leq 93000; \end{aligned}$$

Notice the first two constraints are equivalent to:

$$B = 30C + 60S$$

$$T = 20C + 30S$$

Do you agree with the formulation? If so, you should analyze its solution below:

```

Optimal solution found at step:          0
Objective value:                        0.5455500E+08
Variable          Value          Reduced Cost
   B             7380000.          0.0000000
   T             3990000.          0.0000000
   C              60000.00         0.0000000
   S             93000.00         0.0000000

   Row      Slack or Surplus      Dual Price
   1         0.5455500E+08         1.0000000
   2         0.0000000            5.5000000
   3         0.0000000            3.5000000
   4         0.0000000            235.00000
   5         0.0000000            435.00000

```

There is a hint that the formulation is incorrect because the solution is able to magically produce almost four million Tot sets from only 100,000 sticks.

The mistake that was made is a very common one for newcomers to LP, namely, trying to describe the features of an activity by a constraint. A constraint can always be thought of as a statement that the usage of some item must be less-than-or-equal-to the sources of the item. The last two constraints have this characteristic, but the first two do not.

If one analyzes the dimensions of the components of the first two constraints, one can see there is trouble. The dimensions (or “units”) for the first constraint are:

Term	Units
B	Big sets
$30C$	$30 \text{ [connectors/(Big set)]} \times \text{connectors}$
$60S$	$60 \text{ [sticks/(Big set)]} \times \text{sticks}$

Clearly, they have different units, but if you are adding items together, they must have the same units. It is elementary that you cannot add apples and oranges. The units of all components of a constraint must be the same.

If one first formulates a problem in words and then converts it to the algebraic form in LINGO, one frequently avoids the above kind of error. In words, we wish to:

```

Maximize profit contribution
Subject to:
    Usage of connectors ≤ sources of connectors
    Usage of sticks ≤ sources of sticks

```

Converted to algebraic form in LINGO, it is:

```

MAX = 5.5 * B + 3.5 * T;
      30 * B + 20 * T <= 60000;
      60 * B + 30 * T <= 93000;

```

The units of the components of the constraint $30B + 20T \leq 60,000$ are:

Term	Units
$30B$	$30 [\text{connectors}/(\text{Big set})] \times (\text{Big set}) = 30 \text{ connectors}$
$20T$	$20 [\text{connectors}/(\text{Tot set})] \times (\text{Tot set}) = 20 \text{ connectors}$
60,000	60,000 connectors available

Thus, all the terms have the same units of “connectors”. Solving the problem, we obtain the sensible solution:

Optimal solution found at step:		0
Objective value:		10550.00
Variable	Value	Reduced Cost
B	200.0000	0.0000000
T	2700.000	0.0000000
Row	Slack or Surplus	Dual Price
1	10550.00	1.000000
2	0.0000000	0.1500000
3	0.0000000	0.1666667E-01

4.7 The Nonsimultaneity Error

It must be stressed that all the constraints in an LP formulation apply simultaneously. A combination of activity levels must be found that simultaneously satisfies all the constraints. The constraints do not apply in an either/or fashion, although we might like them to be so interpreted. As an example, suppose we denote by B the batch size for a production run of footwear. A reasonable policy might be, if a production run is made, at least two dozen units should be made. Thus, B will be either zero or some number greater-than-or-equal-to 24. There might be a temptation to state this policy by writing the two constraints:

$$B \leq 0$$

$$B \geq 24.$$

The desire is that exactly one of these constraints be satisfied. If these two constraints are part of an LP formulation, the computer will reject such a formulation with a curt remark to the effect that no feasible solution exists. There is no unique value for B that is simultaneously less-than-or-equal-to zero and greater-than-or-equal-to 24.

If such either/or constraints are important, then one must resort to integer programming. Such formulations will be discussed in a later section.

4.8 Debugging a Model

LINGO has long had a “Debug” command, see Schrage(1989), that may be helpful in finding formulation errors in models that are infeasible or unbounded. Consider the following model.

```

MAX = 2*X + 3*Y;
[CON1] 2*X + Y <= 12;
[CON2] X + Y >= 25;
[CON3] X + 3*Y <= 11;

```

If you try to solve the model, it will be reported as infeasible. Typically, an infeasible or unbounded model contains one or more errors. For an infeasible model, the Debug command (click on: LINGO | Debug) will identify a smallest set of constraints that are infeasible, i.e., cannot all be satisfied. If the infeasibility is due to an error, this set of constraints must contain an error. The Debug report for the above model, if the output level is set to “verbose”, is as follows:

Constraints and bounds that cause an infeasibility:

Sufficient Rows:

(Dropping any sufficient row will make the model feasible.)

[CON2] $X + Y \geq 25$;

Necessary Rows:

(If none of the necessary and sufficient rows are dropped, then the model remains infeasible.)

[CON1] $2 * X + Y \leq 12$;

Necessary Variable Bounds:

(If none of the necessary and sufficient bounds are dropped, then the model remains infeasible.)

$X \geq 0$

The report implies that if you drop the constraint $X + Y \geq 25$, then the model will become feasible. As long as all of the constraints $X + Y \geq 25$, $2 * X + Y \leq 12$, and $X \geq 0$ are retained, the model will remain infeasible. Such a set of constraints is sometimes referred to as an “Irreducible Infeasible Set”, or IIS, for short. For more discussion on infeasibility analysis, see Chinneck(2008). Similar debugging analysis is available for unbounded linear programs.

4.9 Problems

1. The Tiny Timber Company wants to utilize best the wood resources in one of its forest regions. Within this region, there is a sawmill and a plywood mill. Thus, timber can be converted to lumber or plywood.

Producing a marketable mix of 1000 board feet of lumber products requires 1000 board feet of spruce and 4000 board feet of Douglas fir. Producing 1000 square feet of plywood requires 2000 board feet of spruce and 4000 board feet of Douglas fir. This region has available 32,000 board feet of spruce and 72,000 board feet of Douglas fir.

Sales commitments require at least 5000 board feet of lumber and 12,000 square feet of plywood be produced during the planning period. The profit contributions are \$45 per 1000 board feet of lumber products and \$60 per 1000 square feet of plywood. Let L be the amount (in 1000 board feet) of lumber produced and let P be the amount (in 1000 square feet) of plywood produced. Express the problem as a linear programming model.

2. Shmuzzles, Inc., is a struggling toy company that hopes to make it big this year. It makes three fundamental toys: the Shmacrobat, the Shlameleon, and the JigSaw Shmuzzle. Shmuzzles is trying to unload its current inventories through airline in-flight magazines by packaging these three toys in two different size kits, the Dilettante Shmuzzler kit and the Advanced Shmuzzler kit. It's \$29.95 for the Dilettante, whereas the Advanced sells for \$39.95. The compositions of these two kits are:

Dilettante = 6 Shmacrobats plus 10 Shlameleons plus 1 Jig Saw
Advanced = 8 Shmacrobats plus 18 Shlameleons plus 2 Jig Saws

Current inventory levels are: 6,000 Shmacrobats, 15,000 Shlameleons, and 1,500 JigSaws. Formulate a model for helping Shmuzzles, Inc., maximize its profits.

3. A standard problem encountered by many firms when introducing new products is the "phase-out" problem. Given the components for products that are being phased out, the question is: what amounts of the phased out products should be built so as to most profitably use the available inventory. The following illustrates. The R. R. Bean Company produces, packages, and distributes freeze-dried food for the camping and outdoor sportsman market. R. R. Bean is ready to introduce a new line of products based on a new drying technology that produces a higher quality, tastier food. The basic ingredients of the current (about to be discontinued) line are dried fruit, dried meat and dried vegetables. There are two products in the current (to be phased out) line: the "Weekender" and the "ExpeditionPak". In its "close-out" catalog, the selling prices of the two products are \$3.80 and \$7.00 per package, respectively. Handling and shipping costs are \$1.50 per package for each package. It is R. R. Bean's long standing practice to include shipping and handling at no charge. The "Weekender" package consists of 3 ounces of dried fruit, 7 ounces of dried meat, and 2 ounces of dried vegetables. The makeup of the "ExpeditionPak" package is 5 ounces of dried fruit, 18 ounces of dried meat, and 5 ounces of dried vegetables. R. R. Bean would like to deplete, as most profitably as possible, its inventories of "old technology" fruit, meat, and vegetables before introducing the new line. The current inventories are 10,000 ounces, 25,000 ounces, and 12,000 ounces respectively of fruit, meat, and vegetables. The book values of these inventories are \$2000, \$2500, and \$1800. Any leftover inventory will be given to the local animal shelter at no cost or benefit to R. R. Bean. The prices in the catalog are such that R. R. Bean is confident that it can sell all that it makes of the two products. Formulate and solve an LP that

should be useful in telling R.R. Bean how many “Weekender” and “Expedition Pak” packages should be mixed to maximize profits from its current inventories.

4. Quart Industries produces a variety of bottled food products at its various plants. At its Americus plant, it produces two products, peanut butter and apple butter. There are two scarce resources at this plant: packaging capacity and sterilization capacity. Both have a capacity of 40 hours per week. Production of 1000 jars of peanut butter requires 4 hours of sterilizer time and 5 hours of packaging time, whereas it takes 6 hours of sterilizer time and 4 hours of packaging time to produce 1000 jars of apple butter. The profit contributions per 1000 jars for the two products are \$1100 and \$1300, respectively. Apple butter preparation requires a boil-down process best done in batches of at least 5000 jars. Thus, apple butter production during the week should be either 0, or 5000 or more jars. How much should be produced this week of each product?
5. An important skill in model formulation is the ability to enumerate all alternatives. Scott Wilkerson is a scientist-astronaut aboard a seven-day space shuttle mission. In spite of a modest health problem that is aggravated by the zero gravity of space, Scott has been allowed on the mission because of his scientific skills and because a pharmaceutical company has prepared a set of two types of pills for Scott to take each day to alleviate his medical condition. At the beginning of each day Scott is to take exactly one type X pill and exactly one type Y pill. If he deviates from this scheme, it will be life threatening for him and the shuttle will have to be brought down immediately. On the first day of the mission, Scott gets one type X pill out of the X bottle, but in the process of trying to get a pill out of the Y bottle, two come out. He grasps for them immediately with the hand that has the X pill and now he finds he has three pills in his hand. Unfortunately, the X and Y pills are indistinguishable. Both types look exactly like a standard aspirin. There are just enough pills for the full length mission, so none can be discarded. What should Scott do? (Hint: this problem would be inappropriate in the integer programming chapter.)

