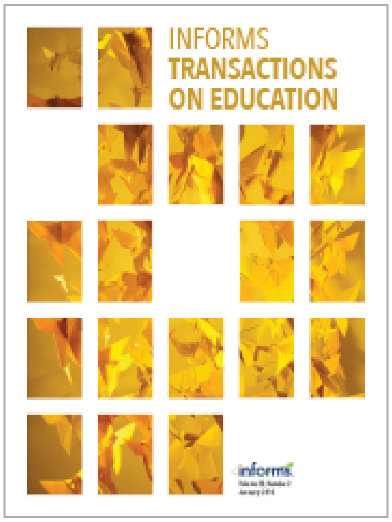

INFORMS Transactions on Education

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Solving “Einstein's Riddle” Using Spreadsheet Optimization

Julian Scott Yeomans,

Solving "Einstein's Riddle" Using Spreadsheet Optimization

Julian Scott Yeomans
Management Science Area
Schulich School of Business
York University
Toronto, ON, M3J 1P3 Canada

syeomans@schulich.yorku.ca

ABSTRACT

A solution to *Einstein's Riddle* is presented using spreadsheet modelling and optimization. Various versions of this problem have been used in introductory management science (MS) classes either as an assignment or as a take-home exam. This riddle has proved to be a challenging problem, since it simultaneously integrates many of the elements that are taught throughout the semester. Namely, the ability to convert a somewhat complicated verbal description into requisite constraints, the creative modelling skills required to transform the problem into an assignment problem-type structure, no "true" or obvious objective function, a difficulty in determining what the (non-obvious) decision variables should be, the use of integer (binary) variables together with either-or constraints requiring satisfaction at equality (an added technical difficulty/challenge), the ubiquitous time issues involved in the solution of integer problems, the numerical representation of numbers by computers that are not readily obvious to business students (i.e. why supposedly integer values may appear in some form of scientific notation) and, most importantly, the ability to appropriately structure the problem formulation into a spreadsheet format for implementation with Solver.

Editor's note: This is a pdf copy of an html document which resides at http://ite.pubs.informs.org/Vol3No2/Yeomans/

INTRODUCTION

Having used numerous variations of the Big Mac problem (Bosch 1993, Erkut 1994) in previous years, I began searching for a suitably challenging replacement problem which could be used to illustrate many of the concepts required in an introductory MS course and one that necessitated sufficiently sophisticated spreadsheet modelling skills. One former student of mine had forwarded an e-mail message to me containing a riddle that he had encountered on the internet. According to this message, "Einstein wrote this riddle last century and said that 98% of the world would not solve it". While this attribution and claim should be considered somewhat dubious at best (see, for example, the comments in http://home.att.net/∼numericana/answer/recreational.htm#einstein5), I will refer to this problem as Einstein's Riddle. The student stated that since I was such a strong proponent of modelling, I should be able to readily demonstrate a spreadsheet optimization approach to this problem. Never wishing to deny anyone the inherent pleasures of spreadsheet modelling, I decided that Einstein's Riddle would actually provide a very testing assignment or take-home examination for my classes. While the final solution may appear apparent, post hoc, the modelling formulation process required to reach it can prove quite difficult and challenging for students.

Einstein's riddle, as it appeared in the original e-mail correspondence, is stated as follows:

Einstein's Riddle

1. There are 5 houses in five different colours.
2. In each house lives a person with a different nationality.
3. These five owners drink a certain type of beverage, smoke a certain brand of cigars, and keep a certain pet.
4. No owners have the same pet, smoke the same brand of cigar, or drink the same beverage.

Required conditions:

-the Norwegian lives in the first house.
-the person living in the centre house drinks milk.
-the owner of the yellow house smokes Dunhill.
-the green house's owner drinks coffee.
-the Dane drinks tea.
-the German smokes Prince.
-the Swede keeps dogs as pets.
-the owner who smokes BlueMaster drinks beer.

  ISSN: 1532-0545

-the person who smokes Pall Mall rears birds.
-the green house is on the left of the white house.
-the British person lives in the red house.
-the person who smokes Blend lives next to the one who keeps cats.
-the person who keeps horses lives next to the person who smokes Dunhill.
-the person who smokes Blend has a neighbour who drinks water.
-the Norwegian lives next to the blue house.

The question is: Who owns the fish?

The Assignment

For the assignment/examination, I provide the students with Einstein’s Riddle exactly as stated above. I require them to formulate a spreadsheet optimization of the riddle with the additional provision that this model must remain linear. They are given a week to construct this model and must submit: a written solution to the problem, all supporting printouts, documentation stating all assumptions used, and an electronic copy of their spreadsheet model.

Teaching Observation:

Confusion can arise with respect to the stated condition that “the green house is on the left of the white house”. This ambiguity occurs when determining whether the green house can appear *anywhere* to the left of the white house, or whether it must be located *immediately* to the left of the white house. If the former interpretation is assumed, then multiple solutions to the riddle can be shown to exist. However, when the latter interpretation is adopted, then the fish ownership solution is unique. The instructor should be aware of this ambiguity and/or clarify exactly which of these interpretations they wish to have adopted in their assignment. The formulation for the riddle in the subsequent section requires that the green house be situated immediately to the left of the white house. The implications of the more “relaxed” interpretation, both for grading purposes and for extending the assignment, will be considered in further detail at the end of the paper.

Teaching Observation:

Because of its nominal link to Einstein, this Riddle has proved to have been a popular puzzle over the years. Due to this enduring popularity, it is possible for students to find solutions to it on the web (see: http://www.mindspring.com/∼mccarthys/puzzle4.htm, http://home.fuse.net/k8dv/riddleans.html, http://www.tpo.net/einstein/answer.htm, http://home.att.net/∼numericana/answer/recreational.htm#einstein5) which makes it imperative for the instructor to require the complete model submission and spreadsheet documentation for the exercise. Additionally, the instructor can eliminate all references to Einstein when distributing the assignment. It should be further noted that all of these puzzle sites have solved the riddle under the assumption that the green house is the immediate neighbour of the white house.

The Solution

Under the assumption that the green house is located immediately to the left of the white house, the German is the only feasible owner of the fish. This solution, including the complete spreadsheet formulation and corresponding Solver instructions, can be found in the Excel file ”Einstein Optimization” (Einstein Optimization.xls) and has been summarized in Table 1.

Table 1:

	House #1	**House #2**	**House #3**	**House #4**	**House #5**
Nationality	Norwegian	Danish	British	German	Swedish
House Colour	Yellow	Blue	Red	Green	White
Drink	Water	Tea	Milk	Coffee	Beer
Pet	Cat	Horse	Bird	Fish	Dog
Cigar	Dunhill	Blend	Pall Mall	Prince	Bluemaster

Mathematical Model and Teaching Observations

Although the riddle, itself, reads in a straightforward fashion, students quickly encounter great difficulty in constructing their spreadsheet optimization models. Over the semester, the course has stressed the need for the explicit identification of the objective(s), decision variables, and constraints during the model formulation phase. Although the riddle states requirements that are quite obviously constraints, a determination of what exactly the objective function and, in particular, the decision variables should be causes considerable consternation. Without the ability to define appropriate decision variables, most students initially flounder about rather aimlessly. This frustration proves very instructive to them, since they begin to appreciate that a systematic approach is requisite in undertaking any significant modelling process even for a small problem such as this; in practice, not all formulations are immediately obvious.

Perhaps the key breakthrough comes with the specification of an effective numbering system for the houses that can be used directly in the constraint calculations. I have found that without this recognition, many students expend an inordinate amount of energy trying to overcome their start-up inertia. This numbering can be accomplished in several ways and in both the spreadsheet and mathematical formulations provided, the houses have been numbered sequentially from 1 to 5, left-to-right (i.e. the leftmost house is number 1, the house immediately to its right is number 2, etc.). Hence, allow subscript i, i = 1, 2, 3, 4, 5, to represent the specific house number.

Teaching Observation:

According to instructor preference, the numbering system may be provided as an additional specification in the assignment instructions. Even with this house numbering scheme, how to actually incorporate it into the solution of the riddle is not readily apparent. However, such an enforced uniformity to the numbering pattern can reduce subsequent grading efforts. On the other hand, it does prove a useful exercise to let students find their own way in formulating their models.

Appropriate decision variables must also be defined, which, as with the house numbering, are not generally obvious to the students. In order to accomplish this task, assignment-type decision variables can be created for the nationality, house colour, drink, pet-type, and cigar-type by defining the following five sets of binary variables, together with their accompanying assignment conditions (see the spreadsheet model, also).

Nationality

Let subscript j, j = 1, 2, 3, 4, 5, represent the nationality of the house occupant with British = 1, Dane = 2, Norwegian = 3, German = 4, and Swedish = 5.

$$Then\ define\ N_{ij} = \begin{cases} 1 & If\ the\ origin\ of\ the\ person\ living\ in\ house\ i\ is\ from\ nation\ j \\ 0 & Otherwise \end{cases}$$

Since only one nationality can be assigned to each house and only one house can be assigned to each nationality, these conditions can be enforced by the following constraints:

$$\sum_{j=1}^{5} N_{ij} = 1,\ i = 1,2,3,4,5\ \ and$$

$$\sum_{i=1}^{5} N_{ij} = 1,\ j = 1,2,3,4,5$$

  ISSN: 1532-0545

House Colour

Let subscript k, k = 1, 2, 3, 4, 5, represent the house colour with red = 1, green = 2, yellow = 3, blue = 4, and white = 5.

$$Define\ H_{ik} = \begin{cases} 1 & If\ house\ colour\ k\ is\ used\ on\ house\ i \\ 0 & Otherwise \end{cases}$$

Since only one colour can be assigned to each house and only one house can be assigned each colour, these conditions can be enforced by the following constraints:

$$\sum_{k=1}^{5} H_{ik} = 1,\ i = 1,2,3,4,5\ \ and$$
$$\sum_{i=1}^{5} H_{ik} = 1,\ k = 1,2,3,4,5$$

Drink

Let subscript l, l = 1, 2, 3, 4, 5, represent the drink with tea = 1, coffee = 2, beer = 3, water = 4, and milk = 5.

$$Define\ D_{il} = \begin{cases} 1 & If\ drink\ l\ is\ consumed\ by\ person\ in\ house\ i \\ 0 & Otherwise \end{cases}$$

Since only one drink can be assigned to each house and only one household can be assigned each drink, these conditions can be enforced by the following constraints:

$$\sum_{l=1}^{5} D_{il} = 1, i = 1,2,3,4,5\ \ and$$
$$\sum_{i=1}^{5} D_{il} = 1, l = 1,2,3,4,5$$

Pet

Let subscript m, m = 1, 2, 3, 4, 5, represent the pet with bird = 1, horse = 2, dog = 3, cat = 4, and fish = 5.

$$Define\ P_{im} = \begin{cases} 1 & If\ pet\ m\ is\ owned\ by\ the\ person\ in\ house\ i \\ 0 & Otherwise \end{cases}$$

Since only one pet can be assigned to each household and only one household can be assigned each pet type, these conditions can be enforced by the following constraints:

$$\sum_{m=1}^{5} P_{im} = 1,\ i = 1,2,3,4,5\ \ and$$
$$\sum_{i=1}^{5} P_{im} = 1,\ m = 1,2,3,4,5$$

Cigar

Let subscript n, n = 1, 2, 3, 4, 5, represent the cigar with Pall Mall = 1, Dunhill = 2, Blue Master = 3, Blend = 4, and Prince = 5.

$$Define\, C_{in} = \begin{cases} 1 & If\ cigar\ brand\ n\ is\ smoked\ by\ the\ person\ in\ house\ i \\ 0 & Otherwise \end{cases}$$

Since only one cigar brand can be smoked in each household and only one household can smoke each cigar brand, these conditions can be enforced by the following constraints:

$$\sum_{n=1}^{5} C_{in} = 1,\ i = 1,2,3,4,5\ \ and$$
$$\sum_{i=1}^{5} C_{in} = 1,\ n = 1,2,3,4,5$$

Teaching Observation:

As with the numbering system, the instructor may wish to enforce a more uniform modelling approach by suggesting the use of these types of variables. In practice, I have not provided such advice since I prefer the students to follow their own solution approach.

Additional Constraint Requirements

Finally, the riddle's constraint can be formulated once a link between the numbering system and the decision variables is recognized. These conditions can be specified in the following manner (see the spreadsheet also):

The Norwegian lives in the first house.

$$\sum_{i=1}^{5} i * N_{i3} = 1 \tag{1}$$

The person living in the middle house drinks milk.

$$\sum_{i=1}^{5} i * D_{i5} = 3 \tag{2}$$

The person living in the yellow house smokes Dunhill.

$$\sum_{i=1}^{5} i * H_{i3} = \sum_{i=1}^{5} i * C_{i2} \tag{3}$$

The person living in green house drinks coffee.

$$\sum_{i=1}^{5} i * H_{i2} = \sum_{i=1}^{5} i * D_{i2} \tag{4}$$

The Dane drinks tea.

$$\sum_{i=1}^{5} i * N_{i2} = \sum_{i=1}^{5} i * D_{i1} \tag{5}$$

The German smokes Prince.

$$\sum_{i=1}^{5} i*N_{i4} = \sum_{i=1}^{5} i*C_{i5} \tag{6}$$

The Swede has a dog.

$$\sum_{i=1}^{5} i*N_{i5} = \sum_{i=1}^{5} i*P_{i3} \tag{7}$$

The beer drinker smokes BlueMaster.

$$\sum_{i=1}^{5} i*D_{i3} = \sum_{i=1}^{5} i*C_{i3} \tag{8}$$

The bird owner smokes Pall Mall.

$$\sum_{i=1}^{5} i*P_{i1} = \sum_{i=1}^{5} i*C_{i1} \tag{9}$$

The green house is situated immediately to the left of the white house.

$$\sum_{i=1}^{5} i*H_{i2} - \sum_{i=1}^{5} i*H_{i5} = -1 \tag{10}$$

The British person lives in the red house.

$$\sum_{i=1}^{5} i*N_{i1} = \sum_{i=1}^{5} i*H_{i1} \tag{11}$$

For the condition that the person who smokes Blend lives next to the owner of cats, there are two possibilities. Under these possibilities, it must be the case that the difference between their house numbers will be either +1 or -1 depending upon whether the Blend smoker lives immediately to the right or to the left of the cat owner. Therefore, this will necessitate the introduction of an "either-or" constraint, together with an appropriately defined binary variable to enforce the requisite condition.

Teaching Observation:

During the semester, students have been introduced to the concept of turning either/or constraints on and off using the Big M method. However, these constraints have all been of the less-than-or-equal-to or greater-than-or-equal-to variety. Thus, adding or subtracting a large number from the RHS is all that is required. This approach does not work when both parts of the either/or condition require equality. Students must determine how this on/off equality can be achieved by a judicious selection of the values and the logic used for the RHS of their constraints; which can prove very challenging for general business students. Since the difference in house numbers must be either +1 or -1, a single constraint format can be adopted that uses the appropriately defined 0/1 binary variable, B, and to have the constraint's RHS evaluate to 1 - (2*B). Such an approach will be employed in this constraint (and all of the other remaining conditions). However, there are numerous alternative formulation approaches that can be employed to enforce these conditions.

Define the binary variable B_1 such that:

$$B_1 = \begin{cases} 0 & If\ the\ Blend\ smoker\ lives\ to\ the\ immediate\ right\ of\ the\ cat\ owner \\ 1 & If\ the\ Blend\ smoker\ lives\ to\ the\ immediate\ left\ of\ the\ cat\ owner \end{cases}$$

Then the requirements for this condition can be implemented by satisfying the condition:

$$\sum_{i=1}^{5} i*C_{i4} - \sum_{i=1}^{5} i*P_{i4} = 1-(2*B_1) \tag{12}$$

The person who keeps horses lives next to the person who smokes Dunhill. As with the previous condition, there are two possibilities and the difference between their house numbers will be either +1 or -1 depending upon whether the horse owner lives immediately to the right or the left of the Dunhill smoker. Therefore, define the binary variable B_2 such that:

$$B_2 = \begin{cases} 0 & If\ the\ horse\ owner\ lives\ to\ the\ immediate\ right\ of\ the\ Dunhill\ smoker \\ 1 & If\ the\ horse\ owner\ lives\ to\ the\ immediate\ left\ of\ the\ Dunhill\ smoker \end{cases}$$

Then the requirements for this condition can be implemented by satisfying the condition:

$$\sum_{i=1}^{5} i * P_{i2} - \sum_{i=1}^{5} i * C_{i2} = 1 - (2 * B_2) \tag{13}$$

The person who smokes Blend has a neighbour who drinks water. Thus, the Blend smoker must live immediately to the right or the left of the water drinker. Define the binary variable B_3 such that:

$$B_3 = \begin{cases} 0 & If\ the\ Blend\ smoker\ lives\ to\ the\ immediate\ right\ of\ the\ water\ drinker \\ 1 & If\ the\ Blend\ smoker\ lives\ to\ the\ immediate\ left\ of\ the\ water\ drinker \end{cases}$$

Then the requirements for this condition can be implemented by satisfying the condition:

$$\sum_{i=1}^{5} i * C_{i4} - \sum_{i=1}^{5} i * D_{i4} = 1 - (2 * B_3) \tag{14}$$

The Norwegian lives next to the blue house. Thus, the Norwegian must live immediately to the right or the left of the blue house. Define the binary variable B_4 such that:

$$B_4 = \begin{cases} 0 & If\ the\ Norwegian\ lives\ to\ the\ immediate\ right\ of\ the\ blue\ house \\ 1 & If\ the\ Norwegian\ lives\ to\ the\ immediate\ left\ of\ the\ blue\ house \end{cases}$$

Then the requirements for this condition can be implemented by satisfying the condition:

$$\sum_{i=1}^{5} i * N_{i3} - \sum_{i=1}^{5} i * H_{i4} = 1 - (2 * B_4) \tag{15}$$

With the decision variables and constraints defined in this way, an arbitrary objective function can be introduced in order to complete the model formulation.

In the spreadsheet optimization, the objective employed is to:

$$\sum_{j=1}^{5} \sum_{i=1}^{5} N_{ij} \tag{16}$$

Teaching Observation:

During my course, I do not instruct the students that Solver can also be used to evaluate a set of equations/constraints for feasibility. Thus, in this problem, they are required to determine that there is no "real" objective for them to maximize or minimize. This provides an added test to their problem solving abilities.

Teaching Observation:

Depending upon the cell formatting used on the spreadsheet, the solutions produced by Solver often include the binary variables expressed in scientific notation; not the integer 0/1 format that the students expect. This "awkward" display of numbers tends to cause an expression of overt skepticism as to the accuracy of spreadsheet solutions by many students. It can therefore be instructive to discuss computer representations of numerical results. This proves especially useful for business students who have not have been previously exposed to such concepts.

Teaching Observation:

One primary purpose of the introductory MS course is to encourage students to effectively and systematically formulate models to help solve disparate problems that may be encountered in the future. The natural tendency for students upon receiving this assignment is to initially solve it "on paper". Certain students who are subsequently unable to formulate an optimization model to the riddle tend to create a "fake" spreadsheet which, they claim, actually "produced" the solution that they found manually. This effort clearly goes against the modelling goals of the class and is the major reason for requesting an electronic copy of their spreadsheet; to see that it does, indeed, create their solution to the riddle. The electronic version can be further checked for authenticity by entering the alternate formulation described below and ensuring that their model can still hold under this other possibility.

Teaching Observation:

To circumvent the above noted difficulty, a challenging instructor might wish to change one of the initial conditions so that no feasible solution is possible. Why would an instructor do such a thing? In order to create a superior pedagogical exercise, this change prevents someone from solving the problem by hand and then submitting a "fake" model which produces this solution. Without the benefit of knowing the solution a *priori*, this change forces the students into actually modelling the problem. For those who complain that they cannot find a solution, I respond that they should be constructing a model of the problem as *stated*. Some time after the assignment has been handed out (and after much effort and frustration on each student's part) I reveal to them that the required condition should be changed back to the true requirement as specified above. If someone has constructed a model using the good formulation techniques stressed throughout the semester, then this new information will require only very minor adjustments to the incorrectly formulated model - only one constraint need be changed. This part of the exercise clearly illustrates that, in practice, not all problems are initially formulated consistently or correctly - sometimes problem specifications contain logical inconsistencies which must be subsequently rectified. However, if models are built in the systematic way stressed throughout the class, then these inconsistencies can be addressed and removed without having to build an entirely new model. (Important teaching note: One should not reveal this inconsistency on the same day as course evaluation forms are filled out).

Teaching Observation:

There is another alternative for assessing whether a student's model does indeed satisfy the problem conditions. As stated, I require *all* students to submit an electronic copy of their optimization model. Since there are myriad approaches to model construction, examining all of the model components can prove extremely cumbersome for very large class sizes. However, if requirement (10) had been interpreted to imply that the green house had to be located *anywhere* to the left of the white house, then it could be re-written as:

$$\sum_{i=1}^{5} i * H_{i2} \leq \sum_{i=1}^{5} i * H_{i5}$$

Under this reformulated condition, there are alternative solutions to the riddle as illustrated in Table 2. Therefore, given this alternate allocation for fish ownership, an instructor can quickly assess the validity of each spreadsheet by changing constraint 10 to an inequality, adding a constraint that assigns the fish to the Dane, and then resolving. If the model is capable of producing this alternate allocation, then this can serve

  ISSN: 1532-0545

to confirm its efficacy. Conversely, if the model fails to produce the alternate fish allocation, then the instructor has immediate proof that it has not been constructed correctly.

Teaching Observation:

An alternate variation on the previous observation is to recognize the ambiguity arising from constraint 10 from the outset and to re-write the assignment to clearly state that the “green house can appear *anywhere* to the left of the white house”. Under these circumstances, an added challenge can be included in the assignment that requires the students to identify all possible solutions to the riddle.

Teaching Observation:

Finally, there exists a perception by some in the OR/MS community that the power of an optimization model is more apparent to students when they cannot solve a problem in any other way. Thus, it might seem that an instructor has needlessly forced a spreadsheet approach to this riddle upon their students. While I do not disagree with this observation in principle, I have found that in practice, even after a semester of instruction, the vast majority of general business students who have only recently been exposed to optimization retain a significant “disconnect” between trivial-sized, introductory problems and large-scale, “real life” problems. I find that it can prove most instructive to provide problems that can be solved using the modelling concepts introduced throughout the class and whose solutions can also be determined (perhaps by brute-force approaches) through some other means. If the students can become sufficiently convinced that a systematic modelling approach proves superior under such circumstances, then they become more strongly convinced that these approaches can be readily applied in much more complex settings. Sometimes this adaptation requires a very slow progression. However, the use of problems such as Einstein’s Riddle can be used as an entertaining catalyst and instructive bridge during this conversion process.

ACKNOWLEDGMENTS

The author would like to express his gratitude to the Editor and Associate Editor for providing numerous, very helpful suggestions and contributions that significantly improved the content and presentation of this problem.

REFERENCES

Bosch, R., 1993, ”Big Mac Attack”, *OR/MS Today*, Vol. 20, No. 4, pp. 30-31.

Erkut, E., 1994, ”Big Mac Attack Revisited”, *OR/MS Today*, Vol. 21, No. 3, pp. 50-52.