Judging Perspective 2014

Lunch Crunch: Can nutritious be affordable and delicious?

Kelly Black
Department of Mathematics, Clarkson University

1 Introduction

The questions participants explored for the 2014 Moody Mega Math Challenge revolved around the Healthy, Hunger-Free Kids act of 2010. The teams were asked to address three tasks: determine caloric needs of an individual, determine the distribution of caloric needs across a large population of people, and determine a way to construct a plan to meet students' caloric needs. Their submissions were to be in the form of a report to the US Department of Agriculture. Judges were asked to evaluate each team's entry based on how they addressed the three tasks.

Every year that I have been involved in this event I have seen remarkable growth in the teams' ability to express and evaluate the models that they construct. This trend continued this year, and the advisors and teachers who worked with the student teams deserve great praise for their efforts. Our hope is that this event provides a goal on which students can focus their energies as well as a context for advisors and teachers to work with students to explore a range of activities associated with the modeling process. When we see the fruits of these efforts we feel joy in the results and are proud of the dedication of the teams and their coaches.

We hope to give you feedback as to how we see the fruits of your efforts, and I will address three aspects of the papers submitted this year. The first topic is the summary, which is the first thing that is read and is the first impression. The second topic discussed includes a wide variety of issues associated with writing. The final topic is an overview of the modeling process itself, including some personal observations from this year's event.

This commentary provides one small snapshot of observations. There are many topics that are not covered, nor are the specifics about the judging activities included. I highly recommend that the commentaries from previous years be reviewed. Many important topics have been purposely omitted here because they have been included in previous commentaries that are available to you.

2 The Summary

Every year we see impressive improvements in the entries that are submitted. This is especially true of the summaries. A good summary should be a single page of text that informs the reader what is included in the report. A technical report is not a mystery, and the only surprise to the reader should be how well written and complete the report is.

The summary should provide three things: context, approach, and specific results. In order for a paper to move through the initial rounds of judging a good summary is a necessity. It is the reader's first impression, and it sends an immediate message to the reader. The reader should understand what the paper is about and why the topic is important. The questions addressed should be clearly stated.

The reader should also have a good idea of the approach used to address the problem. In this year's event, many teams stated the primary source for their model and the basic idea behind the derivation of the model. For example, a team may give the name of the model, such as the Mifflin-St. Joer model, and briefly mention that the idea is to estimate the rate that calories are consumed while at rest.

Finally, the reader should be given explicit results. In this year's event there were three questions. The first question did not result in a single, concrete recommendation, but the second and third questions provided an opportunity for the team to give specific recommendations. It was common to see a summary of those recommendations in the summaries. A summary without such recommendations was at a relative disadvantage compared to other entries.

3 Writing and Style

Reading the narrative is the sole means for us to decide how good a model is. A team may put together a wonderful model and give a tremendous effort, but it does not mean anything if it is difficult to read the result of their hard work. The writing itself should be direct and complete. It should not be a description of the team's activities, but rather it should be a description of the model itself. In this section I will discuss some of the important issues about the writing, including the description of the modeling; using tables and graphs; citations; assumptions; the introduction and conclusions; and the format of the paper itself.

3.A Describing the Model

The first question in the 2014 event is an important question that has attracted a great deal of attention. There are numerous models and approaches that have been developed, and it is appropriate to start with an existing model. It is not appropriate, though, to assume that the reader is familiar with the model or understands the model. It is important to inform the reader of the source, meaning, and limitations of the model.

For the first question, a number of papers simply included a list of the equations and sometimes provided a citation for the model. The resulting narrative is difficult to read, and the model itself can appear to be incomplete without making assumptions about the authors' intention. It is best not to leave such things to the reader's imagination.

It is also important to make sure that proper punctuation and grammar is used with respect to equations. An equation should be treated like an object (a noun). It should be part of a sentence, and punctuation is important.

3.B Tables and Figures

A common way to present information is to use a table or a figure. In this year's event a large number of teams used tables to convey values for parameters or aspects of the models for different populations. Every table and figure should be described in the narrative. The team should guide the reader and let the reader know how to read a table or figure and should let the reader know what features are most important.

This is another area of improvement in this year's event. It was more common to see papers that included a discussion of every table and every figure. Not only are more teams doing a better job of including these descriptions, but the descriptions themselves are getting better.

The teams with the best written entries included more details and more insights into how to interpret the tables and figures. Compared to previous years, it was more common to see papers that have tables and figures that are directly relevant to the point that the team is trying to make. It was also more common to see papers in which the team members provided appropriate discussions dedicated to each table and figure. The teams' advisors are doing a better job of informing and training students in the basics of writing technical reports, and this is one area in which the improvements have been especially noticeable.

3.C Citations and References

Another area that has seen constant improvement is the use of both citations within the text and references at the end of the text. This year it was more common for the papers that were still present in the later rounds to include both citations and references. Some teams made use of footnotes which is acceptable as well. What is important is that the team should make it clear what material they used for an idea or motivation and what part of their document reflects that motivation.

If this trend continues it will be difficult for a team to receive a high rating in the future without using both citations and references. When a paper is read that provides both citations and references the reader is given an implicit message that the writers are conscientious with respect to the background work that they performed. It is also an explicit message that the writers respect and recognize the work of other people.

3.D Assumptions

In last year's Judge's Commentary I said that a bulleted list of assumptions is acceptable, and we understand that under the circumstances of a fourteen hour event it is unreasonable to expect long passages with respect to this aspect of the modeling process. A large number of teams took that advice to heart, and it was quite common to see such lists in the entries this year. Many of the best entries included bulleted lists of assumptions.

Having said that (there is always a "but), I do recommend caution in how those lists are expressed. It was not uncommon to see a section that consisted solely of a list of assumptions. It can be jarring to see a long list with no other text. When a team makes use of bulleted lists they need to take special care to provide transitions and a brief

introduction to smooth the flow of the narrative.

On the whole, though, the increased attention to assumptions is a welcome development. It makes it easier for the reader to understand why certain decisions were made, especially when the team refers back to the list of assumptions later. Additionally, it is becoming increasingly common to see teams that list a set of global assumptions to also include a list of assumptions within each section of their paper that pertain specifically to the information in that section. When a team pays attention to such details it is noticeable and sends a clear message to the reader that the team is thinking through all aspects of the modeling process.

With respect to this year's event, more teams also included additional details about the basic assumptions used in developing their model. Most teams simply included a list of the assumptions. Some of the best papers, however, went beyond this and also included a brief justification for each assumption. This often took the form of one or two sentences for each assumption, and it is a nice touch to provide some motivation for the reasons that a team chose to make use of a given assumption.

3.E Introductions and Conclusions

One noticeable development was that there were a large number of entries that had very brief introductions and conclusions. It is not clear if this was to save space or because of time pressures. Some well written documents and well prepared teams had papers whose conclusions were quite short and did not provide a satisfying end of the paper.

The introduction is admittedly a tricky task. It is common for the introduction to closely mimic the summary, which is perfectly fine considering the time pressures. The conclusion section, though, is something that should not be ignored. It is a bit anti-climactic to read a great paper and then have an abrupt ending. The conclusion is the team's opportunity to acknowledge loose ends, restate conclusions, and tie different aspects of the paper together.

One way to demonstrate the importance of the introduction and conclusions is this: it is not uncommon for a professional to pick up a journal article, read the introduction, read the conclusion, and then decide whether or not to delve deeper into the article. The teams should approach this event with the same attitude about the importance of these two sections. We understand that the time and page limits make this difficult, but the last impression for the reader is the conclusion. This is the last thing read before the judge has to make a decision about how a paper compares to other papers.

3.F Formatting

My personal impression from this year is that the mean length of papers seemed to be longer than in previous years. There were far fewer short papers in the five to ten page range. It is exciting that more teams appear to be expending a greater effort and are making more progress on the problem.

At the same time, there appeared to be more teams that were trying to format the pages in a way to take up more space. For example, there was more white space in the middle of the paper and some teams appeared to be taking creative liberties with the margins. The length of the paper is not important, and there were many excellent papers that did not require the full number of allotted pages. However, when it is obvious that a team is making efforts to fill space it can elicit a negative reaction. It is possible to have a complete paper and stay well under the page limits.

4 Modeling

A broad overview of the modeling for this event is given here. First, an overview of the models that were used is given. Next, a wide variety of modeling issues are discussed including the description of the model, strengths and weaknesses, sensitivity, tables, and notes about regression.

4.A The Models

It was common for teams to make use of an existing model for the first question. This is entirely appropriate and should be encouraged. The first step in the first problem required the students to determine the total energy requirements for an individual. A partial list of some of the more common models can be found below:

- The Harris Benedict FormulaBibliography entryBibliography entry,
- Mifflin St. Joer EquationBibliography entry,
- Estimated Energy Requirement from the Institute of MedicineBibliography entry.

Different modifications of these models were provided by the teams to adjust for a variety of aspects. For example,

some teams adjusted the model to account for different activity levels and some teams adjusted the energy expenditures with respect to the amount of sleep an individual is able to obtain.

The goal for the first problem was to determine the number of required calories for lunch. This was done in a variety of ways. Some teams assumed that students would have three meals with equal calorie intake at each meal. Some teams found different distributions, and other teams adjusted as to whether or not an individual was able to eat breakfast.

When it came to addressing the second and third questions there was a much wider range of approaches. For the second question a common approach was to simply find the mean caloric expenditures for each grade level. More advanced approaches included assumptions about the distribution of weights and heights of students for each grade level. For example, many entries assumed that both the weights and the heights for students are normally distributed, and the teams found a wide variety of ways to estimate the mean and standard deviation. Based on these assumptions, a distribution for the energy expenditures could be inferred.

The third question was the most difficult question to address in a structured, mathematical way. The majority of teams assembled sample menus that met basic criteria and used their approach as an example of the kind of menus that are possible. It was not always clear if they were trying to demonstrate the possible menu choices or trying to demonstrate the methodology that can be used to create a menu.

The third question provided teams with an opportunity to stand out. Creating a structured mathematical model for information that is discrete (menu choices) based on continuous criteria (nutritional requirements, taste, and cost) is an exceedingly difficult task. A number of teams developed algorithms to generate a list of potential menus based on choices from different categories. A small number of teams developed scoring systems for a potential menu and then created a scatter plot with axes based on nutrition versus taste.

4.B Describing the Model

The problem examined in this year's event is an important problem that has garnered the attention of a wide variety of people. Taking advantage of their efforts is the correct way to begin the modeling process. It is still important, though, to provide a brief description of the model and provide some motivation as to why the model was chosen. It was not uncommon to read an entry where an equation was simply stated with no description. Teams that introduced the ideas behind the model, stated the model (including a citation), and then provided a brief description of the different terms in the model were more likely to provide the best impression.

When a team presents the model and can describe the model they are more likely to make a better impression on the reader. Moreover, it helps the reader understand the adjustments and changes made to a model and how it is adapted to the current question. The teams that convey an understanding and context help the reader, and in this case, make it easier for the reader to make the transition required to answer the second question in this year's event.

4.C Strengths and Weaknesses

Modeling is a repetitive process. It generally starts with a simple mathematical description employed as a way to attempt to capture the most basic phenomena of interest. A vital aspect of the modeling process, though, is to go back and make changes and updates as a way to improve the model. This requires that a team step back and ask critical questions about the model.

The short time allowed in this event makes it impossible to take part in this second, vital step in the modeling process. Unfortunately, the student teams are limited in what steps they can take in the process. (It is an opportunity for advisers to go back and ask how to improve on a model and experience the full range of activities.) The teams can, however, take the initial steps in the next stage of the process. By stopping to identify the potential problems and advantages inherent in a model, the team can demonstrate that they understand that there is more to do.

Student teams that provide important insights into the relative advantages and disadvantages of their model are demonstrating that they understand the broader range of activities associated with the modeling process. Given the very short time available to the teams it is expected that the models are relatively simple and should have room for improvement. A team that can provide critical insight into the shortcomings of their model is explicitly demonstrating the next steps to take in their efforts.

4.D Sensitivity

Another important step in the modeling process is a detailed introspection of the implications of a given model. One way to do this is through a sensitivity analysis. The number of teams that address this critical aspect of the modeling process has increased each year, and this year an impressively large number of teams attempted a sensitivity analysis. It is clear that in the future a team will have to make some attempt at this part of the process to be recognized as being part of the highest ranked set of papers.

The down side is that there appears to be some confusion as to what a sensitivity analysis is. The basic idea is that you want to determine what parts of your model will result in the largest change for some small change in the model itself. That is, if you make a small change to the model and then look at the output of your model for a fixed set of inputs, you want to determine which change results in the largest difference in the outputs. This analysis can focus on the parameters in the model, individual terms of the model, or even the model itself.

For example, suppose a team constructs a model that includes the equation

$$y = 3x^2 + 10x.$$

The sensitivity of this model depends on the value of x. If the value of x is close to zero then the linear term results in the greatest change for a given input. The term, $10x$, has the greatest influence, and it is more important that this term be correct. Great attention should be given to insure that the value $10x$reflects the phenomena of interest. On the other hand, if the absolute value of x is larger, then the $3x^2$ term is more important, and it is more important that this term be a correct representation of the phenomena of interest.

As the team moves into the process of refining the model, it may be determined that the $3x^2$term in the model is too sensitive when the values of $x$are large. A subsequent update, for example, may be to change that term to $3x^{1.5}$instead.

4.E Tables

One task that was common for many teams was to construct tables of values. For example, many teams did this to express the distribution associated with energy expenditures for different populations of students. Creating a model that is expressed as a table is a common task at all levels of modeling. There are important considerations, though, when constructing tables and including them in a report.

First, it can be difficult for a reader to understand what a table is the first time it is viewed. The writer must take care in describing the table and making it clear what the table is and how to interpret the table. More importantly, it should be clear how the numbers were calculated. The reader should have enough information to be able to reproduce the table if she wishes to do so.

Additionally, a paper that has a large number of tables can be difficult to read. A team should be careful in how they format and include tables in their paper. The team may want to consider putting the tables at the end of the document or in an appendix while making it clear where the tables are located and how they were generated.

4.F Regression Is Not Modeling. (Redux)

The role of regression and modeling is a topic that comes up every year. This year it was less of an issue. The teams did a better job of using regression in an appropriate manner, and it is encouraging to see improvements in how teams are using this important tool. The advisors' efforts are clear, and they are doing a tremendous job informing students how to use combinations of ideas and techniques in appropriate ways. On a personal note, I believe that the efforts put forth by the advisors in this way are one of the primary benefits of this event, and we owe the advisors a great debt of gratitude.

The primary way in which regression was used this year was to create a polynomial to describe the variation in important parameters. For example, many teams decided that they needed the mean weight as a function of age for a population of people. The majority of those teams were able to find data for specific years over a range of years, and they used the resulting regression curve to interpolate between the given range. It was rare to see a team try to extrapolate using the regression curve, which should always be done with extreme care.

Many teams attempted to find ways to express the variation of a particular variable (generally height, weight, or activity level) using the regression curve and then integrating to find some total number of people. This is an approach that we commonly see each year, but it is important to note that using sums and creating a function that depends on a categorical variable is a perfectly legitimate way to express the results of a model. For example, in this event, a certain aspect of the model might depend on the weight and height of the people, which are continuous variables, but the activity level was often categorized as four discrete values. Some teams tried to use interpolation to express activity levels as a continuous variable, but a number of teams left the activity level as a discrete variable and reacted accordingly. Neither approach is “better” than the other, and I would argue that leaving it as a discrete variable gives the team more time to explore other aspects of their model.

5 Conclusions

The 2014 Moody’s Mega Math Challenge asked students to provide a detailed analysis of the Healthy, Hunger-Free Kids Act of 2010. Three tasks were given to the students asking them to determine caloric requirements for individuals, caloric requirements for a broad population of people, and a way to construct a menu to meet the needs of the

population.

The entries in this year's event continued to impress the readers, and it is getting more difficult to determine and compare the differences between entries. However, there are areas for improvement. In particular, teams should be careful in how they present existing models and equations, and they should be careful about transitions especially with respect to bulleted lists. Teams should also be careful in how they present information in tables and make sure to provide complete descriptions of the tables as well as how they were generated. Finally, teams should pay attention to balance across all sections of their paper especially with respect to the introduction and conclusion.

There has been steady progress in other areas. The most striking improvements were seen in the summaries as well as the overall progress the teams made in addressing all three tasks. Additional improvements were seen, in how the teams progressed in the modeling process. We continue to see more teams paying attention to the analysis of their models, especially with respect to potential weaknesses. Additionally, there are a growing number of teams paying close attention to the issue of a model's sensitivity.

There is little doubt that these improvements are the result of the teams' advisors and teachers. The efforts and dedication of the advisors is an inspiration, and we have been thrilled to see these dramatic improvements over such a short time. The event itself provides a context for our advisors and teachers to make a positive impact, and we are grateful for their efforts.

6 Acknowledgements

Thank you to the folks at the Society of Industrial and Applied Mathematics for their guidance and editorial support in putting this document together. In particular I would like to thank Kathleen LeBlanc and Michelle Montgomery for their help and insights which made this a better document to share with our faculty advisors.

Bibliography

harris-benedict: J. Arthur Harris and Francis G. Benedict, A Biometric Study of Basal Metabolism in Man, 1919

harris-benedict-re-evaluated: A.M. Roza and H.M. Shizgal, The Harris Benedict equation reevaluated, 1984

mifflin-stJoer: Mifflin MD, St. Jeor ST, Hill LA, Scott BJ, Daugherty SA, Koh YO, A new predictive eqn for resting energy expenditure in healthy individuals, 1990

ERR-IOM: George A Brooks, Nancy F Butte, William M Rand, J-P Flatt, B Caballero, How a physical activity recommendation came to be among dietary recom., 2004