

HMMT Scoring System Overview

September 4, 2011

1 Need for a new scoring system

HMMT has gone through many iterations of scoring algorithms, starting with a very basic system including no normalization at all. Throughout the life of the tournament, it has become apparent that such a basic system does not provide sufficiently justifiable comparison between different tests. Here we describe a system that provides a good comparison between the competitors, even with the unusual conditions that have existed in HMMT.

2 Scoring a contest

The eventual goal when scoring a contest is to order the competitors in a ranking list. Additionally, each competitor must be assigned a contribution to their team's score. As a result, our goal is to assign a number to each competitor which represents their relative performance on the contest. Our first priority is to create an accurate method for computing these scores. Our current solution is not as simple as previous methods, but comes with several advantages over alternative solutions.

3 Method

Suppose that we were given the list of numbers corresponding to relative performance on the entire contest. This, combined with knowledge of how difficult each problem is, would allow us to estimate the probability that a given person solved a given problem correctly. Furthermore, it would allow us to estimate the probability of an entire set of problem results. We have assigned prior probabilities for competitor performance and for problem difficulty. After updating these probabilities based on the actual problem results, we choose the most likely set of performances and difficulties. For a more detailed but involved description, see the technical specification.

4 Advantages

Our system has many advantages which we believe outweigh the loss in simplicity in using this scoring system.

- Accuracy for every competitor

One of the problems with our previous system, which involved dividing by the average of the top 10 on each test, was that the problems that determined the score distribution in the top 10 weren't the problems that differentiated competitors in other score brackets, especially when comparing between tests. This system does not give any special attention to any particular score bracket, and competitors at all ranks should see more accurate results.

- Avoids inaccurate problem weights

Previously, our problem czars would assign a predetermined weight to each problem. Although most of the weights were accurate, the human nature of this method inevitably leads to some problems being weighted wrongly, which detracts significantly from the accuracy of the scores. Our new system determines the problem difficulty solely based on the results of the contest with no preconceived notions of difficulty, which could be mistaken.

- Fewer "magic numbers"

The previous method of scoring HMMT involved a very artificial choice of top 10 for the scores to average. Although our system has some parameters which function similarly to the 10 in that they can be tweaked, in this case many choices lead to similar qualitative behavior, and we have chosen numbers for simplicity in these cases. This leads to similar behavior across all skill levels.

- Places the emphasis on doing math

One of our great concerns whenever we consider a scoring system is its ability to be gamed, so to speak. In particular, we are always worried that the optimal strategy for a team is not to solve all the problems that they can. Due to the increased complexity of this scoring system, many people were worried about increased possibility of gaming. In our tests, we found exactly the opposite, that no plausible team strategy would significantly change the results of the contest.

If you would like more detail as to how this system is implemented, see the technical specification. Additionally, our system can be viewed as a slightly modified version of the Rasch model, so many of the qualitative features from the Rasch model carry over.