MBMT Scoring Algorithm

March 30, 2019

1 Statement of Problem

Students at MBMT compete in teams of at most five. Each student takes two subject tests in subjects $\mathsf{s}_1, \mathsf{s}_2 \in \{\text{algebra}, \text{combinatorics}, \text{geometry}, \text{number theory}\}$, which are tests of eight questions of varying difficulty. Also each team takes a team round and a guts round. We will assume that the scores of the team round are fixed regardless of the scoring system.

A scoring system must produce

- For each competitor c taking a subject test s, a subject test score $S_{\mathsf{c},\mathsf{s}}$
- For each competitor c, an individual score I_{c} combining the two subject test scores
- For each team t, a team score T_{t} that incorporates scores from all the tests associated with that team.

There are two main difficulties: combining different subject test scores, and adjusting scores for the team score so that each round is weighted in a consistent manner.

The following procedure is defined for one division only; we run separate computations for the two divisions.

2 Score Shifts

In this section we define a function $\varphi_a(s)$ that shifts scores and will form the basis of our normalization. First, we state the definition:

$$\varphi_a(s) = \frac{e^a s}{(e^a - 1)s + 1}.$$

This is motivated by considering the *sigmoid function*

$$S(x) = \frac{1}{1 + e^{-x}}.$$

This function goes to 0 at $-\infty$ and 1 at ∞, and so is a natural candidate for mapping a score in the interval $[0, 1]$ to some ability $[-\infty, \infty]$. Straightforward computation will now show that $\varphi_a(s) = S(S^{-1}(s) + a)$, so the φ_a shifts the "ability threshold" for a test by a constant amount a.

We will note that this new form of φ_a makes it clear that $\varphi_a \circ \varphi_b = \varphi_{a+b}$, so these shifts are consistent in some sense.

3 Subject Tests

Given a subject test question, its weight is defined to be

$$w = 2 + \ln\left(\frac{N+2}{n+2}\right),$$

where N is the number of people who take the test associated with the question, and n is the number of correct answers. This formula was chosen to satisfy a few properties:

- Convexity — the difference between weights of problems solved by 60% and 65% of people should be less than the difference between weights of problems solved by 5% and 10% of people.
- Relative size — since this year at MBMT we expect roughly 100 people to take each subject test, a problem solved by one person will be worth about 5.5 points, about 2.5 times the lowest possible point value
- (In)sensitivity — the difference between a problem solved by one person and two people should not be drastic; here it is about 0.3 points

The subject test score is simply the sum of the weights of all solved problems. We expect maximum possible subject test scores to be in the range $[20, 30]$.

4 Individual Scores

First of all, to normalize, let M_{s} be the maximum possible score on subject test s, counting unsolved questions. Let $s_{\mathsf{c},\mathsf{s}} = S_{\mathsf{c},\mathsf{s}}/M_{\mathsf{s}}$.

Since some subject tests are harder than others, we want to normalize scores before producing an individual score. In other words, we want to find values a_{s} for subjects s. Then we define

$$I_{\mathsf{c}} = \varphi_{a_{\mathsf{s}_1}}(s_{\mathsf{c},\mathsf{s}_1}) + \varphi_{a_{\mathsf{s}_2}}(s_{\mathsf{c},\mathsf{s}_2}).$$

This way, an individual who solves zero questions receives a score of 0, while a student who solves all question receives a score of 2.

To determine the vector of a's $\mathbf{a}$, we consider the following function:

$$L(\mathbf{a}) = \sum_{\mathsf{c}} s_{\mathsf{c},\mathsf{s}_1} s_{\mathsf{c},\mathsf{s}_2} (\varphi_{a_{\mathsf{s}_1}}(s_{\mathsf{c},\mathsf{s}_1}) - \varphi_{a_{\mathsf{s}_2}}(s_{\mathsf{c},\mathsf{s}_2}))^2.$$

This measures how far apart the normalized scores are for different subjects, weighing higher-scoring individuals higher since the main purpose of this normalization is to compare high-ranking students.

Then, to compute $\mathbf{a}$, we minimize L with respect to $\mathbf{a}$ under the constraint that $\sum_{\mathsf{s}} a_{\mathsf{s}} = 0$. This is to make sure that we don't push all the scores to 0 or 1.

5 Team Scores

This procedure is almost identical to the individual normalization. Each team t has an individual total $i_{\mathsf{t}} = \frac{1}{10}\sum_{\mathsf{c}} I_{\mathsf{c}} \in [0,1]$ as well as normalized team and guts scores $t_{\mathsf{t}}, g_{\mathsf{t}} \in [0,1]$.

In the end, we want to compute

$$T_{\mathsf{t}} = 100 \cdot \left(\frac{1}{2}\varphi_{a_i}(i_{\mathsf{t}}) + \frac{1}{4}\varphi_{a_t}(t_{\mathsf{t}}) + \frac{1}{4}\varphi_{a_g}(g_{\mathsf{t}})\right).$$

(The 100 is for aesthetic purposes.)

We have a similar L function, but there are factors of 2 due to the double-weighting of individuals:

$$\begin{aligned} L(a_i, a_t, a_g) = \sum_{\mathsf{t}} & 2 i_{\mathsf{t}} t_{\mathsf{t}} (\varphi_{a_i}(i_{\mathsf{t}}) - \varphi_{a_t}(t_{\mathsf{t}}))^2 + 2 i_{\mathsf{t}} g_{\mathsf{t}} (\varphi_{a_i}(i_{\mathsf{t}}) - \varphi_{a_g}(g_{\mathsf{t}}))^2 \\ & + t_{\mathsf{t}} g_{\mathsf{t}} (\varphi_{a_t}(t_{\mathsf{t}}) - \varphi_{a_g}(g_{\mathsf{t}}))^2. \end{aligned}$$

We minimize this with the constraint that $2a_i + a_g + a_t = 0$.