ON THE FAIRNESS OF THE GEORGIA OFFICIAL SPELLING BEE RULES

Alexandra L. Wang, Lowndes High School, 1606 Norman Drive, Valdosta, GA 31601
Krishnendu Roy, Valdosta State University, Valdosta, GA 31698

ABSTRACT

We investigate the fairness of the Georgia Official Spelling Bee Rules. This is an important issue, since many students across the State of Georgia participate in such competitions. We perform the simulation optimization method and generate the Spelling Bee system with the Georgia Official Rules. Our simulation outputs show that the Georgia Official Spelling Bee Rules are not fair and are biased due to the random order of seat allocation. We also derive some theoretical results which are consistent with our simulation results.

Keywords
Spelling Bee, probabilistic analysis, simulation modeling, output analysis

INTRODUCTION

Spelling Bee is a competition designed for students at any grade between 4^{th} and 8^{th} to spell English words [4] [6]. In 1925, Nine US newspapers started the National Spelling Bee. When the Scripps Howard News Service took over the sponsorship of the program in1941, the name was changed to the Scripps National Spelling Bee. The competition's main objective is to help students to improve spelling, increase vocabulary, learn concepts, and develop correct English usage [6].

The Georgia Association of Educators (GAE) is the sponsor for the Georgia State Spelling Bee program [4]. This program offers Georgia students an opportunity to exhibit their proficiency in the art of spelling. The local Spelling Bee procedures, rules, and regulations are determined by the GAE.

In the general procedure of the Georgia Spelling Bee program, there are two phases [4]. In Phase I, the final two spelling winners will be determined. In Phase II, the champion spelling winner will be determined. The performance of all spellers is evaluated by the official rules [4]. Because many students across the State of Georgia participate in such competitions, it is important to investigate the fairness of the Georgia Official Spelling Bee Rules.

We study the fairness of the Georgia Official Spelling Bee Rules. Some theoretical results are derived. In probability [1] [7], we show that Georgia Official Spelling Bee Rules are not fair and are biased due to the random order of seat allocation. The current Spelling Bee system does not provide a fair measurement to evaluate all spellers' performance. The evaluation standard is not equally likely.

Due to the high complexity of the Spelling Bee models, theoretical results are difficult to derive. Computer simulation is a powerful tool for solving such real-world problems [2] [3] [5]. We use the simulation based performance evaluation method to model and simulate the Spelling Bee system. The implementation of simulation models is based on the Georgia Official Spelling Rules [4]. Our simulation results are consistent with our theoretical results, which also indicate the unfairness of the official rules.

In sensitivity analysis, our results show that the Spelling Bee system becomes more and more unfair (sensitive) when the difficulty level of spelling words goes up.

According to our study, the Georgia Spelling Bee system is unfair based on its current official rules. Most people will not notice that the system is unfair due to the varying levels of speller's preparation. The degree of how unfair the competition is seems less severe because the better spellers tend to stay in longer. Here, "Better spellers" means their level of preparation is high. Therefore, the Georgia Spelling Bee Official Rules need to be modified. There are other ways to make the Spelling Bee completely fair for all spellers. One feasible way is to adopt the Official Rules from the Scripps National Spelling Bee system.

THE GEORGIA OFFICIAL SPELLING BEE RULES

Georgia Official Rules of the school, system, district, and state level Spelling Bees have been using statewide for so many years. In a general Spelling Bee setting, there are n spellers. Each speller has been assigned a seat randomly with a unique label number from #1 to #n. The spelling order under the Georgia official rules is the ascending order. In each round, the speller with a smaller seat label number goes first.

The Georgia Spelling Bee has two phases. Phase I consists of single elimination. Final two spelling winners will be determined by the following Rule #5 of the Georgia Official Spelling Bee Rules [4]:

"5.When a speller fails to spell a word correctly, he or she must step out of the Bee; and a different word shall be given to the next speller, except as provided for in Rules 13-15 of this section."

In Phase II, the champion spelling winner will be determined by the following official rules #13, #14, and #15 [4]:

"13.When only two spellers remain in the Bee, the Bee procedure changes. When one speller misspells a word, the other speller shall be given an opportunity to spell that same word. If the second speller spells the word correctly, plus the next word on the caller's list correctly, then he or she shall be declared the Winner."

"14.When one of the last two spellers misspells a word and the other, after correcting the error, misspells the new word called to him or her, then the misspelled new word shall be referred to the first speller. If the first speller succeeds in correcting the error in the word and correctly spells the next word on the caller's list, then he or she shall be declared the Winner."

"15.When both spellers misspell the same word, both shall continue in the Bee, and the one who first misspelled the word shall be given a new word to spell. The Bee shall then resume under Rules 13-14."

If a speller is successful at the school system level, then he/she will have the opportunity to advance to the district competition. The district winners and runners-up compete in the state spelling bee with the state winner advancing to the national spelling bee [4].

PROBABILISTIC ANALYSIS

As we discussed in the previous section, there are two phases in the Georgia Spelling Bee implementation. For a particular example in Phase I, if the first n-2 spellers are eliminated in the first round, the last two spellers are automatically being winners and will enter Phase II for further competition. If that is the case, we even do not have a chance to evaluate their performance, whatever their intelligence levels are. Thus we do have a reason to challenge the Georgia Official Rules. Does

every speller have the same chance to be the final two winners in Phase I? Do both final two winners have the same chance to be the champion winner in Phase II?

In order to ensure accurate results, we assume that all spellers have a uniform level of preparation. This means that all spellers have the same probability of spelling a word correctly. Furthermore, this uniform assumption also applies to the probability of spelling a referred word correctly in the next section.

Let *p* be the probability of spelling a word correctly for all spellers and let *q* be the probability of spelling a word incorrectly, which is equivalent to *1-p*. By the model assumption, *p* is the level of preparation. Later on, we also use the term of passing rate for *p*. For the simplicity of the notation, we define p_i be the probability that the i^{th}speller is the first one eliminated in the first round ($1 \le i \le n$). Who has the highest probability of being the first one eliminated in the first round of Phase I? It is a conditional probability. The condition is "be the first eliminated speller in the first round". We derive this probability now.

By definition, the *i*-1 spellers in front of the i^{th}speller (with smaller seat label numbers) have spelled all words correctly and the i^{th}speller is the first one to spell the word incorrectly. Therefore,

$$p_i = p * p * \dots * p * q = p^{i-1}q = p^{i-1}(1-p), \quad for\ 1 \le i \le n.$$

Since $0 < p < 1$, this implies,

$$p_1 > p_2 > p_3 > \cdots > p_n.$$

The probability decreases as the seat label number increases. The speller with the seat label #1 has the highest probability of being the first one eliminated in the first round, and speller with the seat label #n has the least probability. This result disproves the fairness of the Georgia Spelling Bee system. In general, the speller with a greater seat label number has a lower probability of being eliminated in Phase I.

We summarize our results into the following table at different preparation levels for a 5 spellers system.

Table 1: The probabilities of being the first one eliminated

p	p_1	p_2	p_3	p_4	p_5	***Range***
0.10	0.90	0.09	0.009	0.0009	0.00009	0.89991
0.30	0.70	0.21	0.063	0.0189	0.00567	0.69433
0.50	0.50	0.25	0.125	0.0625	0.03125	0.46875
0.70	0.30	0.21	0.147	0.1029	0.07203	0.22797
0.90	0.10	0.09	0.081	0.0729	0.06561	0.03439

At each row, all probabilities are in descending order. The range value decreases as *p* value goes down. This indicates that the Georgia Spelling Bee system is sensitive to lower level of preparation *p*. From a different point of view, the *p* value can be explained as the level of spelling given words. When the difficulty of the words increases, the unfairness of the Georgia Spelling Bee Rules becomes more and more unfair.

Let's consider the last round situation in Phase I now. We are assuming that there are only the top three spellers left. Without loss of generality, we mark them with seat label numbers: #1, #2, and #3. The level of preparation for all three is still *p*. What is their probability of entering finals in the last round? Here, last round means that we have to eliminate only one speller to end Phase I. All probabilities of the remaining part of this section are conditional on the last round assumption. We define p_i be the probability that the i^{th}speller is entering finals in the last round ($1 \le i \le 3$).

For the #1 seat speller, there are two cases of entering finals:

- #1 and #2 spellers are entering finals, or
- #1 and #3 spellers are entering finals.

Therefore

$$p_1 = p * p * (1-p) + p * (1-p) = (1-p)(p+p^2).$$

For the #2 seat speller, there are also two cases of entering finals:

- #1 and #2 spellers are entering finals, or
- #2 and #3 spellers are entering finals.

We have

$$p_2 = p * p * (1-p) \ + (1-p) = (1-p)(1+ \ p^2).$$

Similarly, for the #3 seat speller, the two cases are:

- #1 and #3 spellers are entering finals, or
- #2 and #3 spellers are entering finals.

The probability is

$$p_3 = p * (1-p) + (1-p) = (1-p)(1+p).$$

For a simple algebraic comparison, we have,

$$p_1 \ < \ p_2 < \ p_3.$$

This result also confirms the same seat order advantage: the #3 speller has the highest probability of entering finals while the #1 speller has the lowest. Detailed probabilities at different levels of preparation are listed in the following table.

Table 2: The probabilities of entering finals

p	p_1	p_2	p_3	*Range*
0.10	0.099	0.909	0.990	0.891
0.30	0.273	0.763	0.910	0.637
0.50	0.375	0.625	0.750	0.375
0.70	0.357	0.447	0.510	0.153
0.90	0.171	0.181	0.190	0.019

This table shows that all probabilities of entering finals are increasing as the seat number increases at any levels of preparation. Based on the last column, the current system is very sensitive when the level of preparation is low. Table 2 also shows an important result. When the word level of difficulty is very high, there is a large gap in probability of entering finals between #1 speller and the group of #2 and #3 spellers. From the first row of Table 2, the passing rate for a given word is 10%. The chance of entering finals for #1 is about 10%. However, the chance for #2 is about 91%, and for #3 is about 99%. There is a huge seat advantage for the last two spellers.

SIMULATION MODELS AND OUTPUT ANALYSIS

Due to the complexity of the Phase II model, theoretical results are difficult to derive. We use computer simulation, a powerful tool, to simulate the Georgia Spelling Bee system. Simulation output analysis is performed to evaluate the fairness of the current system.

The overall simulation model contains both phases. We carry all notations and assumptions from previous sections. Java is used here for programming. There are 4 parameters in the input modeling part: the level of preparation *p*, the total number of speller *n*, the total number of simulation iterations *m*, and the re-passing rate p^*. The re-passing rate is only defined for the Phase II part. From the official rule #13, if one speller misspells a word, the other speller shall be given an opportunity to spell that same word. It is reasonable to assume that the "other" speller has a higher chance to pass the same word, because she or he at least can eliminate one wrong spelling combination. This chance or probability is called the re-passing rate. Obviously, it is higher than the normal passing rate *p*.

First of all, we generate the overall Georgia Spelling Bee system. There are 5 spellers in this system. The total number of iterations is 10,000,000. This implies that all 4 digits after decimal point of the simulation output are significant in accuracy [8] [9]. We will perform the simulation output analysis to check if the current system is fair to all spellers. Do they have the equal chance to become the champion at the same level of preparation? The answer is no. The simulation output data is listed in the following table:

Table 3: The probabilities to be the champion

p	p^*	p_1	p_2	p_3	p_4	p_5	*Range*
0.10	0.19	0.0461	0.0473	0.0483	0.3672	0.4911	0.4450
0.30	0.37	0.1048	0.1166	0.1285	0.2524	0.3977	0.2929
0.50	0.55	0.1330	0.1519	0.1741	0.2278	0.3132	0.1802
0.70	0.73	0.1551	0.1724	0.1927	0.2206	0.2592	0.1041
0.90	0.91	0.1823	0.1904	0.1991	0.2088	0.2194	0.0371

We have the following results, which are consistent with the results from the previous section.

- Probability to be the champion increases as the seat number increases.
- The system becomes more and more unfair (sensitive) as the level of preparation goes down. This result is based the range value in the last column of Table 3.
- The last two spellers have huge seat advantage to become the champion.

From row 1 of Table 3, for a close look, the chances of becoming the champion for #1, #2, and #3 spellers are about 4.61%, 4.73%, and 4.83%. But the chances for the last two spellers are 36.72% and 49.11%. The gap is huge and significant.

Furthermore, we simulate the Phase II part of the current system separately. Phase II is designed for the last two winners to compete for the champion, and is determined by official rules #13 - #15. Our simulation output also shows that Phase II is not fair. The output is listed in the following table.

Table 4: The probabilities to be the champion

p	p^*	p_1	p_2	*Range*
0.70	0.75	0.4598	0.5402	0.0804

The #2 speller has the seat advantage too. The probability of the #2 speller to win is about 17% higher than the probability of #1 speller due to the seat order.

CONCLUSIONS

In conclusion, the Georgia Spelling Bee system provides an unfair competition. This system is determined by the Georgia Official Spelling Bee Rules [4]. The chance for all spellers to win the championship is not equally likely even if they have the same level of preparation. The major reason to cause this problem is the random seat assignment. Spellers with higher seat label numbers have a higher chance of becoming the champion. The unfairness of this system becomes more and more serious as the level of spelling words becomes harder and harder. The last two spellers with the largest seat label numbers have a significant (huge) seat advantage (chance) for entering the finals.

Every year, many students across the State of Georgia participate in such competitions. The fairness of the current Georgia Spelling Bee system is an important issue. Therefore, it is imperative that the problem pointed out in this study be fixed.

ACKNOWLEDGEMENT

We would like to thank Mr. Andrew MacLaren of Valdosta State University for his valuable comments, suggestions, and assistance in Java programming.

REFERENCES

[1] Billingsley, P. *Probability and Measure*, 2nd Edition, New York, NY: Wiley, 1986.
[2] Bratley, P., Fox., B.L., and Schrage, L.E. *A Guide to Simulation*. 2nd Edition, New York, NY: Springer-Verlag, 1987.
[3] Fishman, G.S. *Monte Carlo: Concepts, Algorithms, and Applications*. Springer Series in Operations Research, New York, NY: Springer-Verlag, 1996.
[4] Georgia Association of Educators. "2011-2012 State Spelling Bee: Procedures, Rules, and Regulation," GAE, Tucker, GA: http://www.gae2.org/pdf/bee/11-12beebook.pdf, 2011.
[5] Law, A.M. and Kelton, W.D. *Simulation Modeling and Analysis*, 3rd Edition, New York, NY: McGraw Hill, 2000.
[6] National Spelling Bee Official Website. http://www.spellingbee.com, 2012.
[7] Starnes, D.S., Yates, D., and Moore, D.S. *The Practice of Statistics*, 4th Edition, New York, NY: W. H. Freeman, 2010.
[8] Wang, A.L., Xu, Z., and Kicey, C.J. "Asymptotic Statistical Analysis of Pseudo Random Numbers", *Proceedings of the 2011 IEEE International Conference on Computer Science and Automation Engineering*, Piscataway, NJ: IEEE Press , 2011, 542–545.
[9] Wang, A.L., and Kicey, C.J. 2011, "On the Accuracy of Buffon's Needle: A Simulation Output Analysis," *Proceedings of the 49th Annual ACM SE Conference*, New York, NY: ACM Press, 2011, 233–236.