Representing Student Performance with Partial Credit

Yutao Wang, Neil T. Heffernan, Joseph E. Beck
{yutaowang, nth, josephbeck}@wpi.edu
Computer Science Department, Worcester Polytechnic Institute

1 Introduction

The educational data mining community has not been paying much attention to how much assistance a student needs. Feng and Heffernen[1] showed that we can predict student performance better by accounting for amount of assistance, but didn't reduce it to a value that could be shared with students. In this paper we want to see if we can better model student performance by replacing traditional binary measures (correct/ incorrect) with continuous partial credit, which is assigned based on the details in the student responses. The data we use is from ASSISTment, a web-based math tutoring system for 7th-12th grade students. The system helps the student learn the required knowledge by breaking the problem into sub-questions called 'scaffolding' or by giving the student hints on how to solve the question.

2 Partial Credit Algorithm & Evaluations

In the algorithm, students get a maximum score of '1' for responding correctly on the first attempt; otherwise, we assign partial credit by penalizing them for hints, wrong attempts and asking for scaffolding. We use 1 / (# of total_hints) as the penalty for each hint; (# of correct_answers) / (# of total_answers - i) as the penalty for the i^{th} wrong attempt in multiple choice problems; 0.1 as the penalty for each wrong attempt in open response problems and for asking for scaffolding. Table 1 shows an example of two data records, in which two students get the same binary performances but very different partial credit (only some of the features are shown because of limited space).

Table 1. Example of different Partial Credit

	…	#hints	#total hints	#wrong attempts	#correct answers	#total answers	Binary performance	Partial credit
1	…	1	5	0	1	5	**0**	**0.8**
2	…	1	5	2	1	5	**0**	**0.22**

We conducted two analyses on 52,529 data records from 72 students in order to evaluate the predictive power of partial credit in comparison to traditional binary performance.

The first analysis compares the predictive power of the mean value of all the previous performances. As a baseline we used a logistic regression model with binary performance as the dependent, and student id as an independent. We then compare using mean previous partial credit vs. mean previous binary performance as independents in the

model. The results are shown in Table 2. The model with mean previous *partial credit* gets a 0.0019 better R-squared than the one with mean previous *binary credit*. The absolute numerical improvement is small, which indicates that it is difficult to substantially improve student modeling. On the other hand, because knowing mean previous *binary performance* only brings 0.0046 better R-squared over baseline, knowing mean previous *partial credit* brings a relative 41.3% better R-squared compared to it.

Table 2. Comparison between partial credit and binary performance

	R^2 (analysis 1)	accuracy (analysis 1)	R^2 (analysis 2)	accuracy (analysis 2)
baseline	0.1857	0.6846	0.1893	0.6721
binary	0.1903	0.6871	0.1979	0.6757
partial credit	0.1922	0.6884	0.1999	0.6791

In the second analysis, instead of the mean value of previous performances, we compare the predictive power of the trends of these two measures of performance. We choose data in which students have 5 opportunities in one skill and use the previous 4 performances to create a trend line to predict the value of the 5^{th} performance. To smooth the result and generate a bounded prediction between 0 and 1, we also apply a logistic regression model to this prediction. The result is shown in the fourth and fifth column of Table 2. Using partial credit improves R-squared by about 0.002 compared to using binary performance. This result is consistent with analysis 1.

3 Conclusions and Future work

In this paper we present a naïve algorithm to assign partial credit given detailed student responses. Partial credit performance contains much more information than binary performance, which is currently used in almost all researchers in the educational data mining field. Evaluations show that partial credit improves student model fitting by only a small absolute value but a high relative value compared to the binary performance.

One question we are interested in is how to refine the algorithm to better fit student data and infer student knowledge. Also, we are interested in finding out in which situations partial credit does better than binary performance, so we can use it as an efficient complement to all the current models that use only the binary performance.

References

[1] Feng, M., Heffernan, N. T. Can We Get Better Assessment From A Tutoring System Compared to Traditional Paper Testing? Can We Have Our Cake (Better Assessment) and Eat It too (Student Learning During the Test)? *Educational Data Mining,* 2010. Pittsburgh, PA.