How to Take Into Account a Student’s Degree of Certainty When Evaluating the Test Results

Joe Lorkowski\textsuperscript{1}, Olga Kosheleva\textsuperscript{2}, and Vladik Kreinovich\textsuperscript{3}

\textsuperscript{1}Department of Computer Science
\textsuperscript{2}Department of Teacher Education
University of Texas at El Paso
500 W. University
El Paso, Texas 79968, USA
lorkowski@computer.org, olgak@utep.edu, vladik@utep.edu
1. Need to Take into Account the Student’s Degree of Certainty

- Student’s grade on a test depends on whether his/her answers are correct.
- However, this approach does not take into account how confident the students is in his/her answer.
- In real life, it is not so bad if a person realizes that his knowledge is weak: he/she may consult another expert.
- The situation is much worse if a decision maker is absolutely confident in his/her wrong decision.
- From this viewpoint, when gauging the student’s level of knowledge, it is desirable:
  - to explicitly ask the student how confident he/she is in the corresponding answer, and
  - to take this degree of confidence into account when evaluating the resulting grade.
2. What We Do in This Talk

• Some tests already solicit the confidence degrees from the students, and take them into account.

• However, the existing techniques for taking these degrees into account are semi-heuristic.

• It is therefore desirable to come up with well-justified techniques.

• To come with such techniques, we use decision making theory.

• We emulate a situation when the decision is made by a group of specialists, including the current student.

• We estimate the expected contribution of this student’s knowledge to the quality of the resulting decision.
3. Simple Case: Description

- According to decision theory, decisions made by a rational agent can be described as follows:
  - to each possible situation, we assign a numerical value called its *utility*, and
  - we select an action for which the expected value of utility is the largest possible.

- Let us start with the simplest case, where we only have one question with two possible answers: $A_1$ and $A_2$.

- Let $P_1$ be the student’s degree of confidence in $A_1$; then, the degree of confidence in $A_2$ is $P_2 = 1 - P_1$.

- Let’s assume that for each answer $A_i$, we know the optimal action.

- For example, for each of the two medical diagnoses, we know an optimal treatment.
4. Simple Case: Analysis

- Let $u_{i,j}$ be the utility corresponding to the case when
  - the actual situation is $A_i$ and
  - we use the action which is optimal for $A_j$.

- The fact that the action corresponding to $A_1$ is optimal for $A_1$ means that $u_{1,1} > u_{1,2}$; similarly, $u_{2,2} > u_{2,1}$.

- Let us suppose that we know the probabilities $p_1$ and $p_2 = 1 - p_1$ of both situations.

- Then, we select $A_1$-action if its expected utility is larger, i.e., if
  $$p_1 \cdot u_{1,1} + (1 - p_1) \cdot u_{2,1} \geq p_1 \cdot u_{1,2} + (1 - p_1) \cdot u_{2,2}.$$

- So, we select $A_1$ if our degree of confidence $p_1$ in $A_1$ exceeds the threshold:
  $$p_1 \geq t \overset{\text{def}}{=} \frac{u_{2,2} - u_{2,1}}{(u_{1,1} - u_{1,2}) + (u_{2,2} - u_{2,1})}.$$
5. How to Estimate the Probabilities of Different Alternatives under Expert Uncertainty

- Let us denote the number of experts by \( n \).
- Let us assume that for each expert \( k \), we know \( k \)'s degree of confidence \( p_{1,k} \) in \( A_1 \); then \( p_{2,k} = 1 - p_{1,k} \).
- In general, we do not have prior reasons to believe that some experts are more knowledgeable than others.
- So we assume that all \( n \) experts are equally probable to be right: \( P(k\text{-th expert is right}) = \frac{1}{n} \).
- Thus, \( p_1 = \text{Prob}(A_1 \text{ is the actual alternative}) = \sum_{k=1}^{n} P(k\text{-th expert is right}) \cdot p(A_1 \mid k\text{-th expert is right}) \).
- So, \( p_1 = \frac{1}{n} \cdot \sum_{k=1}^{n} p_{1,k} \); similarly, \( p_2 = \frac{1}{n} \cdot \sum_{k=1}^{n} p_{2,k} \).
6. How to Estimate the Student’s Contribution to the Correct Decision

• Once we add the student as a new expert, with \( p_{1,n+1} = P_1 \) and \( p_{2,n+1} = P_2 \), then \( p_1 \) changes to:

\[
p'_1 = \frac{1}{n+1} \cdot \sum_{k=1}^{n+1} p_{1,k} = \frac{n}{n+1} \cdot p_1 + \frac{1}{n+1} \cdot P_1.
\]

• We should gauge the student’s contribution by the expected utility \( \bar{u} \) caused by the change in \( p_1 \).

• The value \( \bar{u} \) is proportional to \( p_+ - p_- \), where:

\[
p_+ \text{ is the prob. that we switched to the correct decision},
p_- \text{ is the prob. that we switched to the wrong decision}.
\]

• \( p_+ \) is the probability that \( p_1 < t \) but \( p'_1 \geq t \), i.e., that

\[
p_1 \in \left[ t + \frac{1}{n} \cdot t - \frac{1}{n} \cdot P_1, t \right].
\]
7. How to Estimate the Student’s Contribution to the Correct Decision (cont-d)

- For large \( n \), the interval \( \left[ t + \frac{1}{n} \cdot t - \frac{1}{n} \cdot P_1, t \right] \) is narrow, so \( p_+ \approx \rho(t) \cdot \left( \frac{1}{n} \cdot P_1 - \frac{1}{n} \cdot t \right) \) is linear in \( P_1 \).

- Similarly, \( p_- \) is linear in \( P_1 \), hence \( \bar{u} \) in linear in \( P_1 \).

- So, the grade should be a linear function of \( P_1 \).

- We assign \( N \) points to a fully confident correct answer (\( P_1 = 1 \)) and 0 to fully confident incorrect one (\( P_1 = 0 \)),

- Then, for \( P_1 \in (0, 1) \), we assign:
  - \( N \cdot P_1 \) points if \( A_1 \) is the correct answer, and
  - \( N \cdot P_2 \) points if \( A_2 \) is the correct answer.
8. Discussion

- We wanted to distinguish between:
  - the bad situation when a student is absolutely confident in the wrong answer \( (p_1 = 0, p_2 = 1) \), and
  - a not-so-bad situation when a student is ignorant but understands his/her ignorance: \( p_1 = p_2 = 0.5 \).

- In the first situation, we assign 0 points.

- In the second situation, we assign \( N \cdot 0.5 \) points.

- So, someone with no knowledge can get 50%.

- This means that we need to appropriately change the thresholds for A, B, and C grades.
9. Case of $s > 2$ Possible Answers $A_1, \ldots, A_s$

- A student assigns, to each $A_i$, his/her degree of certainty $P_i$, so that $\sum_{i=1}^{s} P_i = 1$.

- We select the answer $A_{i_0}$ with max expected utility:

$$\left( \sum_{j=1}^{s} p_j \cdot u_{i_0,j} \right) - \left( \sum_{j=1}^{s} p_j \cdot u_{i,j} \right) \geq 0 \text{ for all } i.$$

- Here too, the expected change in utility is linear in $P_i$:

$$\bar{u} = u_0 + \sum_{i=1}^{s} a_i \cdot P_i.$$

- Usually, there are no a priori reasons why one incorrect answer is better than another incorrect answer.
10. Case of $s > 2$ Possible Answers (cont-d)

- So, we assume that the utilities $a_i$ corresponding to each incorrect answer $A_i$ are the same $a_i = a$; then:
\[
\overline{u} = u_0 + a_c \cdot P_c + a \cdot \sum_{i \neq c} P_i = u_0 + a_c \cdot P_c + a \cdot (1 - P_c).
\]

- The utility is a linear function of the student’s degree of confidence $P_c$ in the correct answer.

- Hence, the grade should be linear in $P_c$, i.e., $N \cdot P_c$.

- So, we give the student $N \cdot P_c$ points, where $P_c$ is the student’s degree of confidence in the correct answer.
11. **How Do We Combine Grades Corresponding to Different Problems?**

- Our grades were proportional to expected utilities.
- Usually, different questions on the test are independent from each other.
- It is known that in this case, the overall utility is equal to the sum of the corresponding utilities.
- Thus, the overall grade for the test should be equal to the sum of the grades of individual questions.
- Hence, we arrive at the following recommendation.
12. Resulting Recommendation

• Let us consider a test with \( T \) questions
  
  \[ q_1, \ldots, q_t, \ldots, q_T. \]

• For each question \( q_t \), a student is given several possible answers \( A_{t,1}, A_{t,2}, \ldots \)

• For each question \( q_t \), let \( c(t) \) be the correct answer and let \( N_t \) be the number of points for this answer.

• For each possible answer \( A_{t,i} \), student provides his/her degree of confidence \( P_{t,i} \) that this answer is correct:
  
  \[ P_{t,1} + P_{t,2} + \ldots = 1. \]

• Then, the recommended student’s grade is
  
  \[ g = \sum_{t=1}^{T} P_{t,c(t)} \cdot N_t. \]
13. Acknowledgment

- This work was supported in part by the National Science Foundation grants:
  - HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and
  - DUE-0926721.
- The authors are very thankful to Dr. Michelle D. Miller for her very interesting talk that inspired this research.