

Computation of Assessing the Knowledge in One Domain by using Cognitive Skills Levels

Rania Aboalela
Computer Science
Kent State University
Kent, USA

Javed Khan
Computer Science
Kent State University
Kent, USA

ABSTRACT

Large number of questions are required to assess the knowing of the concepts in one domain. Novel methods are presented to map between the concepts in the domain and between the concepts and the questions of the assessment based on the cognitive skill prerequisite relation. The verbs of Bloom's Taxonomy are used to determine the cognitive skills levels. The cognitive skill levels refer to the levels such as if a learner has acquired the state at the level of understanding, or applying, or analyzing, or evaluating, or creating a concept. The state of achieving knowing or not knowing a certain concept state at certain skill level is called, concept state. Previously, three types of concept states proposed known as Verified Skills (VS), Derived Skills (DS) and Potential Skills (PS). VS and DS refers to the concept state that the concept is learned or not learned by the learner. PS refers to the concept state that the concept is ready to be learned or not ready to be learned. The experiment proved that using the cognitive relation between the concepts increases the amount of the estimated concepts, even though the number of tested concepts may be minimized and eliminated under the conditions laid down by the target cognitive skill levels. Also, the researchers compute the probability of knowing the concepts of two evaluated learners by using Bayes' Theorem.

General Terms

Programming, Algorithms, Knowledge Assessment, Bloom Taxonomy, Pedagogical, Computation.

Keywords

Assessment of Knowledge; Algorithm, Cognitive Skills, Concept Space.

1. INTRODUCTION

This research introduces the cognitive relation methods between the concepts in one domain. In this research, the concentrating was in the domain of the applied since such as Computer Science and Software Engineering. The domain of applied science needs to identify the internal cognitive relation between each concept in the domain to find out an accurate assessment of the knowledge in such domain. The assessment of knowledge in one domain is introduced by Falmagne Cosyn, Doignon, & Thiery [1]. They concentrate in assessing the knowledge in the domain such as Mathematical Science. The cognitive difficulty variations of learning was not identified. In this research, an internal cognitive relation between the concept was identified. The original taxonomy was created in the research work [2]. The Blooms' Taxonomy arranges what the learner has to be learned in a hierarchy of six levels. In 2001, the six major categories were changed from noun to verb forms and renamed [3]. In this research, the verbs of the revised Blooms' Taxonomy to identify the prerequisite cognitive relation between the concepts in the

domain were used. The six verbs are inferred by skill number that indicates the cognitive difficulty as: "1" means recall, "2" means understanding, "3" means applying, "4" means analyzing, "5" means evaluating and "6" means creating. In addition, the research work used an intelligent internal relation between the concepts to assess the knowledge in one domain such as [5]. They identify an ontological relation between the concepts in the course and test questions. They introduced Topic Dependency Graph (TDG). Some research works combining the revised Bloom's Taxonomy with Knowledge Space such as that [6-8]. In these works, they automatically discovering and extracting the Bloom's Taxonomy from the text in one knowledge space. [8] visualized the concept space of course content by three dimensions: Syllabus, Ontological and Cognitive dimension. Moreover, the research work of [9,10] validate the efficiency and the importance of using the parameter of conative skill level to assess the knowledge in one domain.

2. THE METHODS OF THE COGNITIVE RELATIONS ZONES OF THE CONCEPTS IN THE ASSESSMENTS

The researchers propose making the link between any two concepts a verb of the skill which needs to be learned was proposed, for example: to "apply" a concept B, the individual must know the prerequisite concept A at skill level 2, which is the understanding level. The apply is indicated by the number 3 in the link.

2.1 The Zone of Verified Skills

Verified Skills is defined as where there is a direct evidence that a student knows particular concept C_x at a particular cognitive skills level k it is considered to be a part of verified set $VS(k)$. To illustrate the VS, let's consider gathering evidence by a question Q , which can ascertain that a student knows a specific concept C_x . The verified skills satisfy the condition that:

If $(Q_i, C_x)Lk$ & C_x is correct answer $\rightarrow C_x \in VS(k) \forall C_x \in$ completely correct answer concepts.

Where, $Q_i \in$ Test questions, $C_x \in$ Tested concepts, $Lk \in$ Bloom link of level k , $VS(k) \in$ Verified skills at level K and $(Q_i, C_x)Lk$ means existing link between the question Q_i and the concept C_x at level K . The link means that to answer Q_i correctly C_x must be learned at level k . Figure 1 shows verified skill link.



Fig 1: Verified Skills (VS) relation

2.2 The Zone of the Derived Skills at Level 2, DS (k=2)

Derived skill is defined as where there is indirect evidence that a student knows a particular concept C_i at a cognitive skills level 2 (the understand level) it is considered to be a part of DS (2). In other words, if there is indirect evidence that the concept C_i is understood by the student then it will belong to DS ($K=2$). The condition of the relation is expressed as the following:

If C_i is not a member in a verified set but there exists two links such that $(Q_i, C_x)L_k, (C_i, C_x)L_m$ & $C_x \in VS(k)$ that it is a member in VS and $m=2$ and $k \geq m$, then C_i is a member in DS at level 2, i.e. $C_i \in DS(2), Q_i \in$ Test Questions. $C_x \in VS, C_i \in$ another concept in the concept space. The $(Q_i, C_x)L_k, (Q_i, C_x)L_m$ means existing link between the Question Q_i and the Concept C_x at level k and m respectively. Figure 2 illustrates DS relation at level 2.

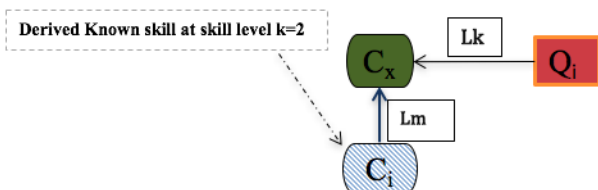


Fig 2: Derived Skill (DS) relation at level 2

2.3 The Zones of Derived Skill, DS (k>2)

Definition: DS ($k>2$) means that there is a direct evidence a student knows a particular concept C_y at a cognitive skill level 2, and there is indirect evidence he knows it at a cognitive skill level higher than a cognitive skill level 2. In other words, it would tell that by inference a student could either apply/analyze/evaluate/create, a particular concept C_y . The relation condition is illustrated as the following: If C_y is known i.e. it is in DS(2) or VS(2), and if all level k support nodes of C_y i.e. $S(C_y, k)$ is in $VS(2) \cup DS(2)$, then C_y will be considered as a Derived Skill at level k . in other words. If $C_y \in DS(2) \cup VS(2)$ and $S(C_y, k)$ is subset of $DS(2) \cup VS(2) \rightarrow C_y \in DS(k)$. Figure 3 illustrates DS ($k>2$).

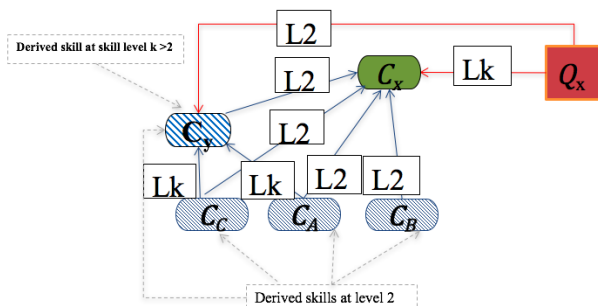


Fig 3: Derived Skills relation at level >2

2.4 The Zone of Potential Skill PS [(k>2)]

Definition: Potential Skill ($k>2$) is defined as where there is indirect evidence that a student knows a particular concepts A at a cognitive skills level higher than 2 (apply/ analyze/

evaluate/ create), it is a part of PS ($k>2$). The relation condition is illustrated as the following:

Let $S(A, k)$ is the support set of A at level k . If every node in the $S(A, k)$ is subset of $VS \cup DS$ at any level (doesn't matter because the purpose is only to guarantee that the set is known) i.e. $S(A, k) \subset VS() \cup DS()$, but there is no evidence that A is known, then A is in potential skill set PS(k) i.e. $A \in PS(k)$, where $C_d, C_x \in (VS)$ and $C_c, C_a, C_b \in (DS)$ and $L_k \in$ Bloom's link at level k . Figure 4 illustrates Potential Skills relation.

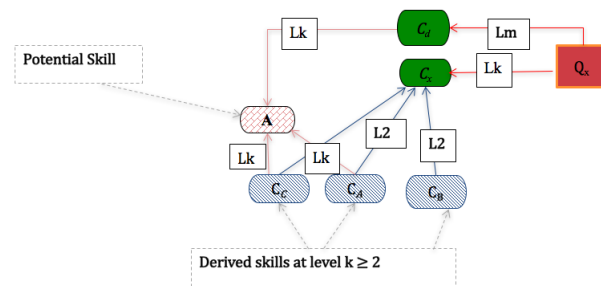


Fig 4: Potential Skills relation

3. CONCEPT MAPPED TESTING AND EVALUATION METHOD

To measure the student learning setting up a concept mapped testing and evaluation method was considered. A test is composed of set of questions. The learners are required to answer the questions based on their knowledge. Grader evaluates the student knowledge based on the answers. In conventional evaluation, a grader grades the answers and assign a quantitative score for the student. This research slightly modifies the evaluation method where the grader instead of a numerical score, is asked to evaluate if there is evidence in the answer that the student has succeed or failed to attain learn a concept at a certain cognitive skill level. This modified method is called concept mapped testing & evaluation method. Each tested concept in the assessment domain is labeled to the question based on the cognitive level as the following theory: To answer the question Q_i correctly the concept C_x must be known at skill level k . Figure 5 shows the relation



Fig 5: The theory of connection the question to the tested concept

4. THE PROBABILISTIC METHODS FOR ESTIMATION THE LEARNING STATES OF THE CONCEPTS IN EACH ZONE

The existence of the concepts in a complex domain such as there are variant prerequisite relations for the concept and the concept could be inferred by more than one concepts. Also, the conflicted evaluations of the concept such as the learner has contradiction in the answer of the questions asked same concept. It would there are many evidences or references infers about the target skills of the concepts in the learning states zones. Sometimes it would be many questions asked about the concepts in the domain and it would be a contradiction between the answers of them. Also, it would be errors in the estimation of the learning states. In real exams, there are other phenomenal like lucky guess or careless

mistakes. People have variant levels of initial knowledge. Accordingly, the probability computation should be used. This fact was studied in [11].

The Bayes' Theorem is used to conclude an accurate probability of the learning states of the assessed concepts. The Bayes' Theorem

$$P(C_j^{Lk}|R) = \frac{P(R|C_j^{Lk}) * P(C_j^{Lk})}{P(R|C_j^{Lk}) * P(C_j^{Lk}) + P(R|\bar{C}_j^{Lk}) * P(\bar{C}_j^{Lk})}$$

Where,

- C_j^{Lk} denotes knowing the concept C_j at skill level Lk
- \bar{C}_j^{Lk} denotes not knowing the concept C_j at skill level Lk
- $P(C_j^{Lk})$ is the unconditional probability of knowing the concept C_j^{Lk} , which is the initial probability of knowing the concept C_j^{Lk} . It is the rate of the correct responses to the questions asked about the concept C_j^{Lk} .

$$P(C_j^{Lk}) = \frac{\text{Total number of the correct answers to the questions asked about the concept } C_j^{Lk}}{\text{Total number of the questions asked about the concept } C_j^{Lk}}$$

- $P(\bar{C}_j^{Lk})$ is the unconditional probability of not knowing the concept C_j^{Lk} , which is the initial probability of knowing the concept C_j^{Lk} . It is just the rate of the incorrect responses to the questions asked about the concept C_j^{Lk} .

$$P(\bar{C}_j^{Lk}) = \frac{\text{Total number of the incorrect answers to the questions asked about the concept } C_j^{Lk}}{\text{Total number of the questions asked about the concept } C_j^{Lk}}$$

- $(R|C_j^{Lk})$ is the conditional probability of the event that the responses in R occur, conditional independence on the event of knowing the concept C_j^{Lk}

- R is the set of responses.

$$P(R|C_j^{Lk}) = \prod_{r=1}^n P(Q_r|C_j^{Lk})$$

= The multiplication of the response data given knowing the concept C_j^{Lk} .

Where, Q_r is response to the question qr . "r" is an integer number that indicates index of the questions.

$$\prod_{r=1}^n P(R|C_j^{Lk}) = \{P(Q_1|C_j^{Lk}) * P(Q_2|C_j^{Lk}), * \dots * P(Q_n|C_j^{Lk})\}$$

$P(Q_r|\bar{C}_j^{Lk}) = g$, when there is dependency between Q_r and C_j

$$P(\bar{Q}_r|C_j^{Lk}) = m$$

$$P(Q_r|C_j^{Lk}) = 1-e$$

$$P(\bar{Q}_r|\bar{C}_j^{Lk}) = 1-e$$

- \bar{Q} is incorrect response to the question r

- Q is correct response to the question r

"g" error value such as lucky guess. "m" error value such as and mistakes, and "e" indicates any type of error. The error values assumed based on the type of question, is it direct or indirect question asked about the skill level of the concept C_j

5. THE EXPERIMENT

The researchers organized an experiment to prove the efficiency of identifying the relations of the cognitive skill level between the concepts to maximize the estimation of measurement the concepts from few tested concepts.

5.1 The Experiment Setup

A human subject test is organized to prove the efficiency of the methods. The test composed of 9 questions, which are selected from the midterm questions that have been given to the learners by the instructor of the class. The class is CS 61002 Algorithms and Programming in the Computer Science department. The test was introduced online in one session. The participants are 154 graduate students, attending the class. In this setup, the questions are specially redesigned to directly test a certain skill level of each concept belonging to the assessment domain. The 9 questions are asked about 18 concepts at certain skill levels. The 18 tested concepts are the concepts in the set of VS. Table 1 shows the number of the tested concepts at each cognitive skill level. The concepts domain is analyzed to find out the states of the concepts set of the derived skills and potential skills. The concepts in the zones DS and PS have never been tested but they would be estimated from the result of the questions asked about the tested concepts set of VS

Table 1. The number of tested concepts at each skill level

Skill Level L	Number of Tested Concepts (Verified Skills)
L2	7
L3	4
L4	2
L5	1
L6	4
Sum	18

5.2 The Efficiency of the Methods

This experiment proves that using the proposed methods associated with the cognitive relations optimizes the knowledge assessment. The result of the evaluation of the perfect learner shows that the amount of the estimated knowledge of the assessed assessment which uses just numerical methods. The perfect learner's answers are used to calculate the experiment footprint and the size of footprint of each relation method. Table 2 and Figure 6 show the size of footprint according to the perfect learner. The perfect learner is the student who gives a correct answer to all the asked questions. As evident the size of VS footprint is 18, the size of DS footprint is 31 and the size of PS footprint is 31, which means that if the learner answered the questions correctly, then it would tell that he knows certain levels of each of these 18 concepts in addition to the estimated 31 concepts by using DS method, and he is ready to know an additional 31 concepts in certain levels of each of them, even though he was tested particularly for 18 skills. By combining the presented knowledge assessment methods with the cognitive relation, the researchers can maximize the amount of the estimation knowledge of the assessed learners.

Table 2. Size of Footprint

Skill Level	Verified VS	Derived DS	Potential PS
L2	7	12	13
L3	4	11	11
L4	2	3	2
L5	1	2	2
L6	4	3	3
Sum	18	31	31

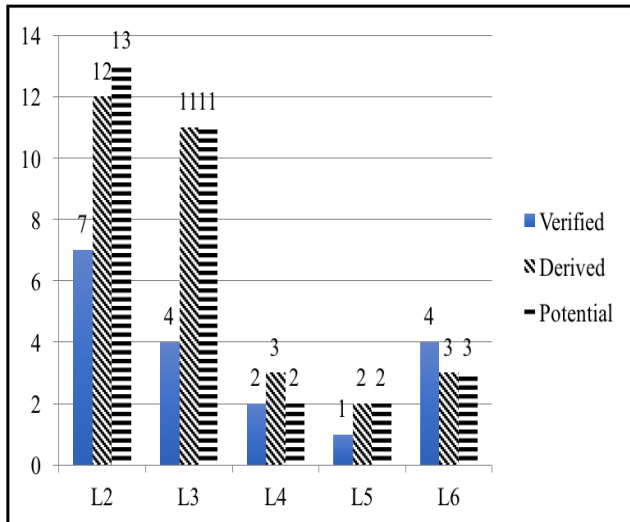


Fig 6: The Size of footprint

5.3 The Probability of the Concept States Based on Human Subject Test

To find an accurate result of knowing or not knowing the concepts set in the concept states zones, the direct questions are designed to asked directly about the skills levels of the estimated concepts in DS and PS.

This is 38 extra questions are designed especially for testing the accuracy of the estimated probability of knowing and not knowing the concepts. An error values are given to the responses to the questions based on the type of the question. If the question is directly asked about the concept the error value $e = 0.1$. If the question is indirectly asked about the skill of the concept the error value $e = 0.2$. Also, a portion of 0.03 is added to the error value of the multiple-choice questions.

The advantage of the existing information of the estimated knowledge and the real knowledge was taken to show the accuracy of the probability of knowing the concepts. The probability of knowing the concepts on a condition of several existing data about the concept is important to know the exact probability of the true concept state of the assessed learner. Figure 7 and figure. 8 show the probabilities according to the estimation of DS Zones of knowing the concepts of the perfect learner and selecting the laziest learner St# 23 (who didn't answer all the questions correctly), respectively. Similarly, a comparison of the probability of knowing the concepts based on PS was made. Figures 9 and 10 show the accuracy of the probabilities according to the estimation of PS Zone. Figure 9 shows the comparison of the evaluation of the perfect learner, whereas Fig.10 shows the comparison of the evaluation of selecting the laziest learner St# 23 (who didn't answer all the questions correctly).

In the figures 7, 8, 9 and 10, each concept is presented by three columns. The first column indicates the estimated

probability of knowing the concept, the second column indicates the probability of real response to the question directly asked about the concept; whereas the third column indicates the probability based on the probability information of the first and second column by using Bayes' Theorem. The x axis represent the concept index number and Y represent the probability value.

As proposed in the analysis of DS and PS, the estimated probability of knowing the concept in DS is inferred based on the probability of the tested supported concept. In this investigation, the probability of errors values in the probability of knowing the estimated concept was increased arbitrary by the investigators. In the direct question which is asked about the same estimated concept the errors values are less. Also, an error value to the type of the question, whether it is multiple choice or an open-ended question, was assigned. If the question type is multiple choice and asks indirectly about the concept, then the probability of error is higher than the open-ended question which asks directly about the concept. From this point, little difference was observed between the estimated probability and the probability of a real response to the direct test on the same concept, even though both indicate the same result of knowing the concept.

Regarding the VS method, the comparison between the estimation of knowing the concept by method and the direct question is identical skills of the concepts which are tested. In other words, the concept at a certain skill is tested in the both two types of the organized questions, indirect question, which not asked about the skill level but it asked about related skill of the prerequisite concept. indirect question called open question (OQ) and direct question, which asked about the skill level directly DQ. The second question of DQ is to confirm the first response to the question of indirect question type. Fig. 11 and fig.12 show the comparison between the evaluation based on the first responses and the evaluation based on the first responses and the computation of the two responses. The researchers assigned the same probability of the error value to both types of the questions, OQ and DQ, was assigned. Therefore, there is no difference between the evaluation if the learner gives the same response either correct or incorrect response in the two questions. If he/she gives incorrect response the third column which refers to the response evaluation will disappear in the graph, since the value is 0

6. CONCLUSION

This paper proposed novel methods to find the zones of the prerequisite relations between the concepts associated with the cognitive skills levels. The efficiency of the methods to minimize the number of tested concepts upon the cognitive skill prerequisite relation is proved. The computation of the probability of the concept states based on human subject test to find an accurate estimation of knowing the concepts is proposed. The connection between the concepts and their presentation in graph view provides optimistic material to improve the adaptive assessment in the most important replacements and acceptance tests such as TOEFL, GRE and GMAT. The proposed analysis methods to realize the connected concepts could decrease the number of tested concepts and, consequently, the assessment time. The most benefit in these areas is that the assessment could provide the exact concepts known or not known by the assessed individual, and they can see their result in graphs rather than just numerically.

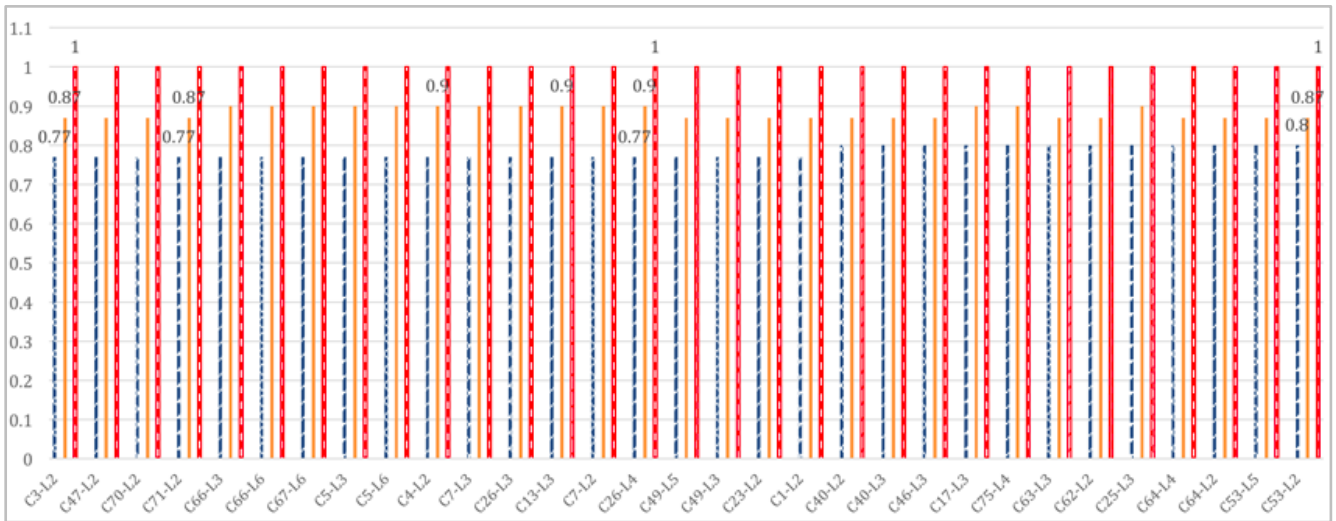


Fig 7: The Comparison Between the Probability of Knowing the Concepts Based on DS Method, Real Responses and Computed Probability of Perfect Learner

X: Concept # at the considered skill level

Y: The probability of knowing the concept at the considered skill level

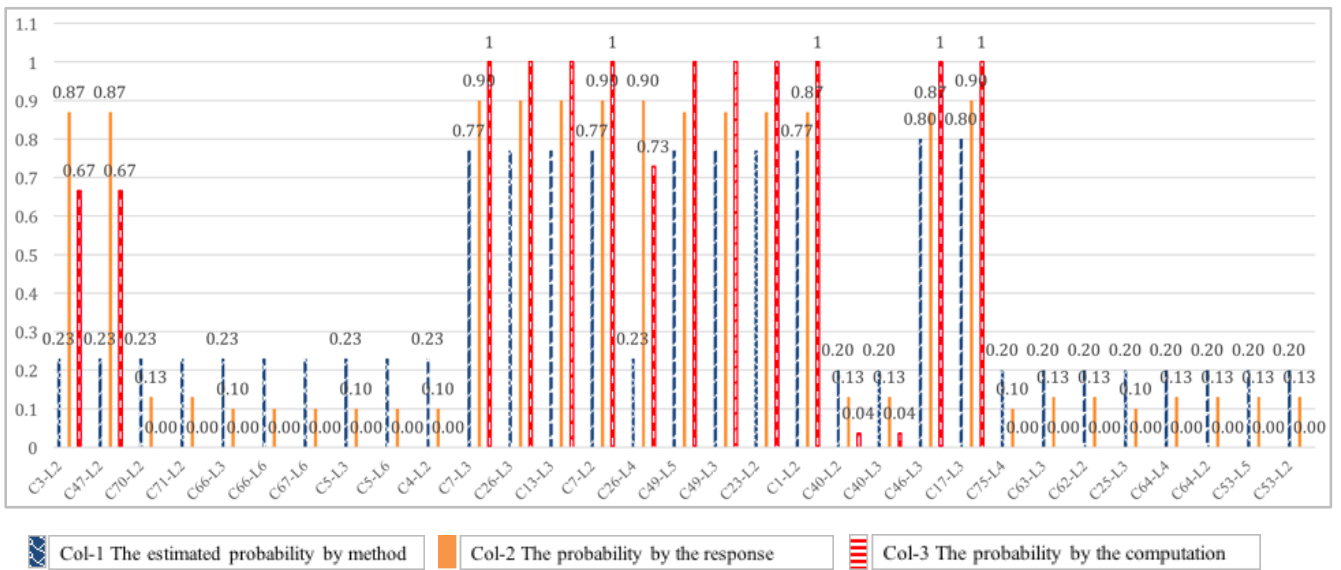


Fig 8: The Comparison Between the Probability of Knowing the Concepts Based on DS Method, Real Responses and Computed Probability of laziest Learner (Student# 23)

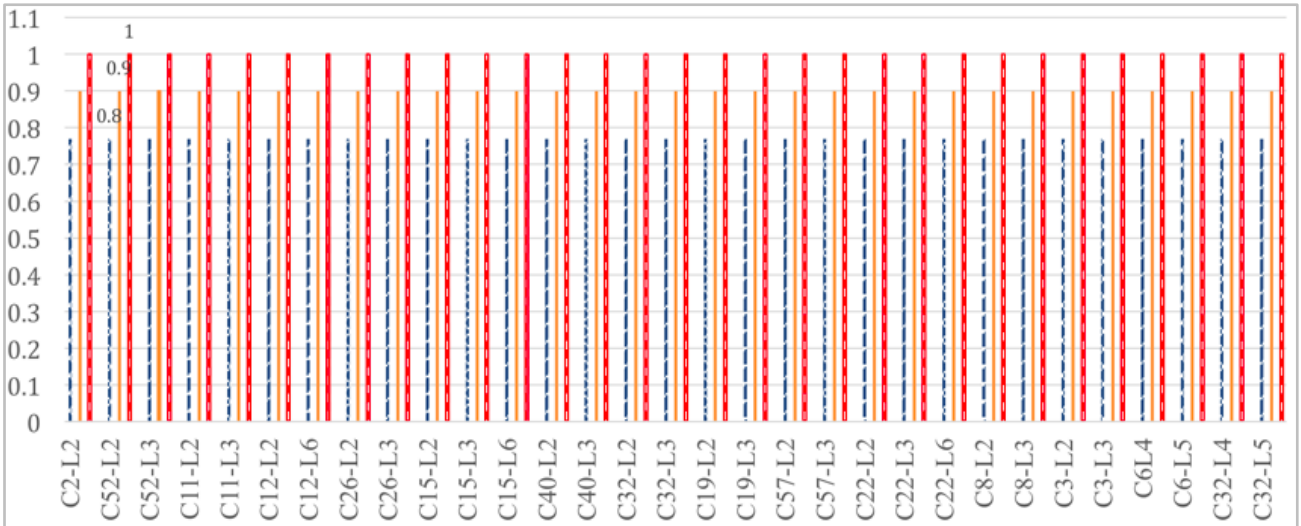


Fig 9: The Comparison Between the Probability of Knowing the Concepts Based on PS Method, Real Responses and Computed Probability of the Perfect Learner

X: Concept # at the considered skill level

Y: The probability of knowing the concept at the considered skill level

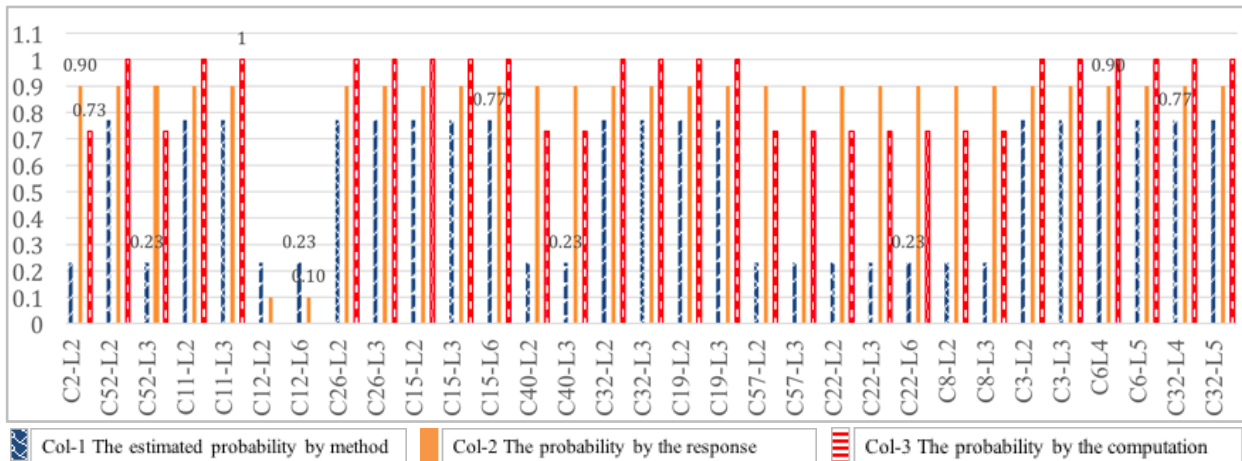


Fig 10: The Comparison Between the Probability of Knowing the Concepts Based on PS Method, Real Responses and Computed Probability by the Laziest Learner

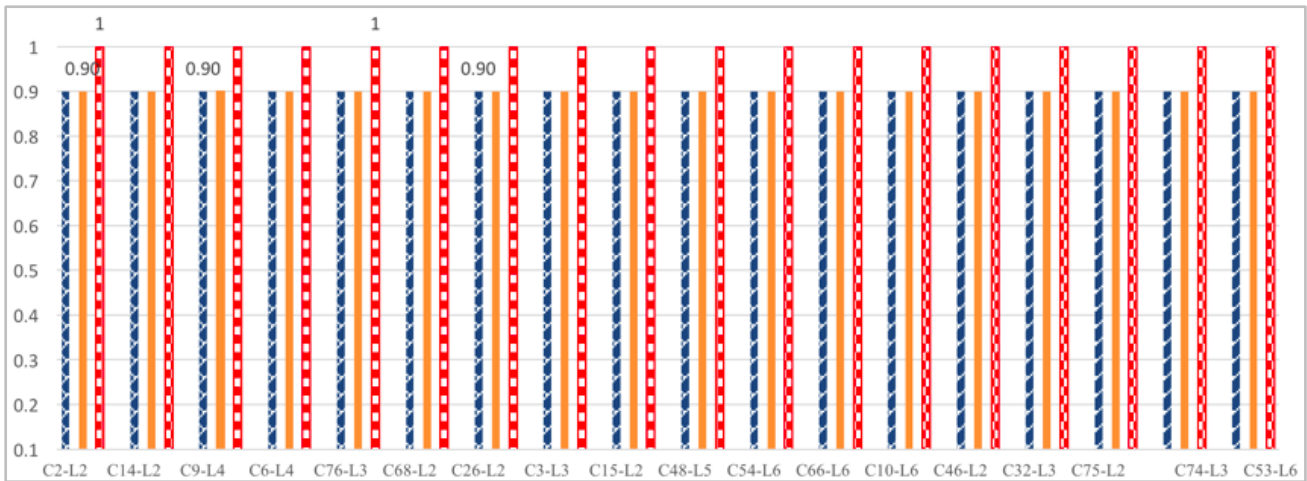


Fig 11: The Comparison Between the Probability of Knowing the Concepts Based on VS Method, Real Responses and Computed Probability of Perfect Learner

X: Concept # at the considered skill level

Y: The probability of knowing the concept at the considered skill level

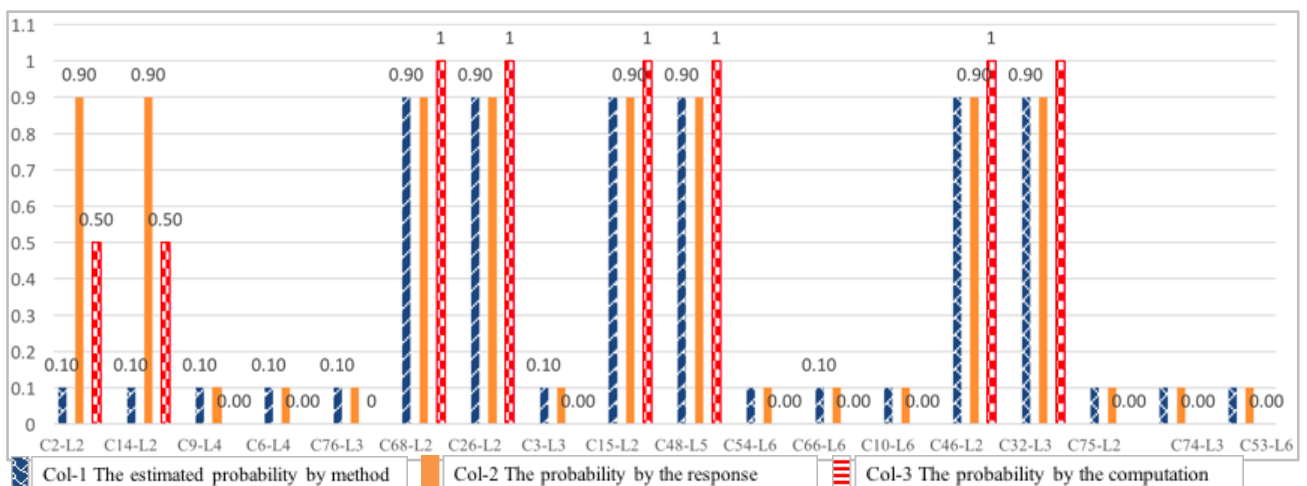


Fig 12: The Comparison Between the Probability of Knowing the Concepts Based on VS Method, Real Responses and Computed Probability of Laziest Learner

7. REFERENCES

- [1] Falmagne, J.-C., Cosyn, E., Doignon, J.-P. and Thiery, N., 2003, The assessment of knowledge in theory and in practice. In Proceedings of the International Conference on Integration of Knowledge Intensive Multi-Agent Systems, pp. 609-615.
- [2] Bloom, B.S., Engelhart, M.D, Furst, E.J., Hill, W.H., and Krathwohl, D.R. (Eds.). 1956. Taxonomy of Educational Objectives. The Classification of Educational Goals, Handbook I: Cognitive Domain. David McKay Company, New York.
- [3] Anderson, L.W., Krathwohl, D.R., Airasian, P. W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., et al., 2001. A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York, U.S.A, Longman.
- [4] Khan, J., Hardas, M., Ma, Y., 2005. A Study of problem difficulty evaluation for Semantic Network Ontology based intelligent courseware sharing. In proceeding of IEEE/WIC/ACM the International Conference on Web Intelligence, WEB Intelligence, pp. 426-429.
- [5] Nafa, F. and Khan, J. Conceptualize the Domain Knowledge Space in the Light of Cognitive Skills. In Proceedings of the 7th International Conference on Computer Supported Education. SCITEPRESS-Science and Technology Publications, pp. 285- 295, 2015.
- [6] Nafa, F., Khan, J., Othman, S., and Babour, A. 2016. Discovering Bloom Taxonomic Relationships between Knowledge Units Using Semantic Graph Triangularity Mining, In proceeding of the International Conference on

- Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), Chengdu: IEEE, pp. 224-233.
- [7] Nafa, F., Khan, J., Othman, S. and Babour, A. 2016. Mining Cognitive Skills Levels of Knowledge Units in Text Using Graph Tringularity Mining. In the proceedings of Web Intelligence Workshops (WIW), International Conference (IEEE/WIC/ACM), IEEE, pp. 1-4.
- [8] Aboalela, R. and Khan, J. 2015. Visualizing concept space of course content. In the proceeding of IEEE 7th International Conference on Engineering Education (ICEED). Japan, November 2015. pp. 609-615 doi: 10.1109/iceed.2015.7451512
- [9] Aboalela, R. and Khan, J. 2016. Model of Learning Assessment to Measure Student Learning: Inferring of Concept State of Cognitive Skill Level in Concept Space. In proceeding of the 3rd International Conference on Soft Computing & Machine Intelligence (ISCMI), Dubai, United Arab Emirates, pp. 189-195. doi: 10.1109/ISCMI.2016.26
- [10] Aboalela, R. and Khan, J. 2016. Are we asking the right questions to grade our students in a knowledge-state space analysis? In proceedings of IEEE 8th International Conference on Technology for Education (T4E), Mumbai, pp. 144-147.
- [11] Aboalela, R. A. 2017 An Assessment of Knowledge by Pedagogical Computation on Cognitive Level Mapped Concept Graphs. Doctoral Thesis. OhioLink ETD. kent1496941747313396. Kent State University.