AN ASSESSMENT OF KNOWLEDGE BY PEDAGOGICAL COMPUTATION ON COGNITIVE LEVEL MAPPED CONCEPT GRAPHS

A dissertation submitted

to Kent State University in partial

fulfillment of the requirements for the

degree of Doctor of Philosophy

by

Rania Anwar Aboalela

Graduation August 2017

Dissertation written by

Rania A. Aboalela

M.S., King Abdulaziz University

B.S., King Abdulaziz University

Approved by

______________________________________ , Chair, Doctoral Dissertation Committee

______________________________________ , Members, Doctoral Dissertation Committee

____

____

____

Accepted by

______________________________________ , Chair, Department of Computer Science

______________________________________ , Dean, College of Arts and Sciences

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

At the start of any new venture, it is my practice to take a moment to give thanks. I'd like first to thank almighty God for granting me good health, ability and opportunity to accomplish my goals. I thank my adviser Dr. Javed Khan for challenging me to push myself harder to fulfill expectations, and for supporting me in my endeavors. Along with my advisor, I am also grateful for having an exceptional doctoral committee in Dr. Arvind Bansal, Dr. Austin Melton, and Dr. Omar De LA Cruz Cabrera. I also have gratitude for professor Dmitry Ryabogin for serving in my committee, and for professor Mahbobeh Vezvaei for her thorough editing of my dissertation both conceptually & mathematically. I would also add special thanks for Hessah Alqahtani for standing with me.

My special thanks to the staff of Department of Computer Science, especially for Janet Kotila, Marcy Curtiss, Nathan Thomas and Brenda Boykin for their smiles and kind words.

I would also like to take a moment to appreciate the support of my parents and my family, especially my grandmother Aziza Al Kateeb, my lovely aunts Ghada Al Himali and Ebtisam Al Himali, and my soul mother Shadia Al Himali. Also, I thank the Ministry of Higher Education of Saudi Arabia for their financial support for my graduate study.

Last, I extend many thanks to my colleagues and friends, especially Maha Allouzi and Amal Babour and Amany Alnahdi. I add my gratitude for the other unnamed people who generously supported me with prayer and good wishes.

Rania Anwar Aboalela

18 August 2017 Kent, Ohio

CHAPTER 1

Introduction

Learner assessment is an integral part of the learning process. The two assessment techniques used for the assessment of a topic, which have gained popularity in science education over the past decade, are: Concept Mapping (CM) (Novak, 2008), and Bloom's Taxonomy (BT) (Anderson, et al., 2001). However, despite the existence of tools and techniques concerned with assessing a topic such as CM and BT, educators continue to use the archaic method of assigning numerical quantity percentage values to measure learner learning. This research study presents an alternate, quality-centric assessment method using Cognitive Skill Levels. The study aims to change the approach used in assessment methods from quantity-centric to quality-centric evaluation. The determination of the need for change in approach is based on the observation that a combination of the two techniques, CM and BT, is the best way to ameliorate current assessment methods. The researcher for this study has set up an intelligent combination of two aspects: the concept-mapped knowledge domain graph of the knowledge assessment's resources and the analysis, algorithms, and evaluation methods. When designing a graph for learning assessment, efforts were made to design an appropriate, applicable and simple Concept-Mapped Knowledge Domain that could improve and simplify the learner-learning assessment process. Based on this first step, a theory is then proposed, connecting the assessment and the assessed domain.

A study on intelligent design attempting to measure knowledge using graphs (Khan, Hardas, & Ma, A Study of Problem Difficulty Evaluation for Semantic Network Ontology Based Intelligent Courseware Sharing, 2005) investigated the difficulty in evaluation by using Semantic Network Ontology-based Intelligent Courseware Sharing. This study inspires the use of thoughtful design in the creation of a new measurement scale. For designing a graph, a model of Concept Space is proposed, named Cognitive Level Mapped Concept Graph (CLMCG). CLMCG is implemented using a scheme-based analysis of mapped graph and logical inference, that in turn provides the assessment of learner knowledge in terms of graphs inspired by Khan, Hardas, & Ma (2005). They introduced conceptual graphs of the ontological relation between the concepts. We modified the graphs by applying Bloom's Taxonomy to identify the internal relationship between the concepts in the knowledge domain. The process of CLMCG maps the entire body of concepts in one space, and it is useful for measuring learners' knowledge in Concept Space. It can also provide details on the practical analysis of conceptual quality-centric learning, in addition to quantity-centric learning. Concept Space is the zone of the related concept in the assessed domain. From the aspect of the analysis algorithms and evaluation methods, an Assessment Theory of Cognitive Skills in Concept Space is proposed, namely TCS^2. The Assessment Theory TCS^2, is a theory (a coherent group of tested general propositions that can be used for explanation and prediction for a class of phenomena) that helps us objectively (algorithmically) understand and assess the learning states and skill levels of a learner. It is applicable to the conceptual content, the concepts themselves and their relationships which define a specific knowledge domain. The Cognitive Skill Levels used in the study are based on Bloom's Taxonomy (Anderson, et al., 2001). The Assessment Theory TCS^2 involves four kinds of methods: (1) The mapping methods used to create the CLMCG; (2) the concept-mapped testing and evaluation method; (3) the theories to produce the sets of Concept States, and, (4) the TCS^2 Analytics, which identify the process to estimate the sets of the Concept States at Cognitive Skills Levels. In the TCS^2 theory, Cognitive Skill Levels are employed to provide learner knowledge as a precise concept set that can be formulated and analyzed. Six basic Concept States are proposed in this Theory: Verified Known Skills set (VKS), Derived Known Skills set (DKS), Potential Known Skills Set (PKS), Verified Known Unknown Skills set (VNS) and Derived Known Unknown Skills set (DNS) and Potential Known Unknown Skills Set (PKS). The VKS consists of the concepts at skills level known by evidence, DKS consists of skills known by inference, PKS consists of skills that are ready to be known by inference, VNS consists of concepts at skills level that are not ready to be known by evidence, DNS consists of concepts at skills level that are unknown by inference and PKS consists of concepts at skills level that are not ready to be known by inference. For the fourth component of the proposed TCS^2 theory, an Assessment Analysis process will detect the exact skills considered in the Concept Space. The result of the assessment analysis method estimates learner knowledge about the concept based on the six Concept States: VKS, DKS, PKS, VNS, DNS, and PNS. Each assessed learner would get his/her result as sets of six Concept States with accurate probability. The probability measurements are based on Bayes' Theorem. The researcher conducted two experiments to both validate the proposed TCS^2

theory and estimate values of the tested theories. The TCS^2 theory is validated by computing the match between the estimated skills sets VKS, DKS, PKS, VNS, DNS, and PNS, and the actual tested concept sets. The high accuracy of the matches validates the presented TCS^2 theory.

1.1 Analysis Algorithms and Methods (The Assessment Theory TCS^2)

The TCS^2 theory involves four kinds of methods: (1) The mapping methods to create the CLMCG; (2) the concept-mapped testing and evaluation methods; (3) the analysis methods to produce the sets of Concept States; and, (4) the TCS^2 Assessment Analytics, which identify the process to estimate the sets of the Concept States at Cognitive Skills Levels.

1.1.1 The Cognitive Level Mapped Concept Graph (CLMCG)

CLMCG is a concept-mapped knowledge domain that consists of an organized knowledge domain (such as a computer operating system, data structure, etc.) in a matrix structure with three relational dimensions: syllabus, ontology, and cognitive skills. The syllabus dimension is the occurrence of the concepts in a formal textbook format, in that it retains the chapter, section, sub-section, etc. The ontological dimension links the concepts in terms of class, part, and instance. The cognitive skill dimension is the relationship between concepts in terms of the prerequisite knowledge needed to transition from one concept to another, in order to attain a particular level of skill in the evaluated concepts.

1.1.2 Concept-mapped testing and evaluation method

To measure learner learning, we set up a concept-based testing and evaluation method. In conventional testing, a learner is given a set of questions, which he/she answers. A grader then evaluates the responses and assigns a quantitative score to the learner. The evaluation method in this study was slightly modified, in that the grader was also asked to evaluate whether there is evidence in the response as to whether the learner has succeeded or failed to attain a certain Cognitive Skill Level. This is termed as concept-mapped evaluation or grading. In this setup, the questions can also be specially designed to directly measure a certain skill level for certain concepts. This process is called Direct Concept-Mapped Testing. The direct questions evaluate the same skill level as designed by the instructor, while directly addressing the level of the concept in the learner. Two types of questions are thus offered: (1) open questions and (2) direct questions. Open questions are open-ended questions without restrictions. An open question is a conventional question prepared by the instructor and could implicitly test the skill level. The direct questions are specially designed to address a certain skill level of the tested concept.

1.1.3 Analysis Methods to Produce the Sets of Concept States

The Assessment Analysis Method is useful to identify the Concept States in the assessment and connect those concepts with the existing CLMCG. This analysis method maps the concepts in the assessment with the questions, using links to indicate the skill levels as verbs that connect the assessed concepts in the domain. Consequently, the concepts at the skill levels presented in the assessment connect with the existing CLMCG and produce the zones of the Concept States. Specifically, the revised Bloom's Taxonomy level is manipulated to estimate a learner's Concept States by using analysis methods from the TCS^2 theory. The theory identifies the assessment result as Concept States (Verified, Derived, Potential, and Disabilities) to ascertain the exact skill levels of concepts learned by the learners.

1.2 Potential Theory & Research Questions

1. The Verified Known Skill set VKS(K)
2. The Derived Known Skill set DKS (K=2)
3. Support Node (SN) and Support Set (SS)
4. The Derived Known Skill set DKS (K>2)
5. Potential Known Skill set PKS
6. The Verified Known Unknown Skill set VNS(K)
7. The Derived Known Unknown kill set DNS (K=2)
8. The Derived Known Unknown Skill set DNS (K>2)
9. Potential Known Unknown Skill set PNS

1.3 Dissertation/Research Objective:

The objective of this research study was to ascertain whether a model can be constructed that can help objectively (algorithmically) understand and assess the learning states and skill levels of a learner, with respect to the content and concepts (and their relationship) in a specific knowledge domain. Armed with the latest developments in graph representation techniques, computer/automated inference technology, and pedagogical theories, this dissertation provides a framework towards achieving this goal.

1.4 Contributions of this Dissertation

The main contribution of this dissertation is to present novel analysis methods of knowledge assessment, which assess all the concepts in one space. The graph paradigm of a semantic/ontological scheme simplifies the difficulty involved in the question-design process. It also provides precise measurement in knowledge assessment by focusing on organizing all the concepts in one space. The method may be considered more precise and efficient due to the use of the inference algorithm, which is faster and uses minimum testing to extract maximum information. Identifying the cognitive link between the existing concepts in one domain increases the accuracy of the estimation of the assessed concepts. Moreover, the number of estimated concepts increases, even though the number of tested concepts may be minimized and eliminated under the conditions laid down by the targeted skill levels. The components of the methods contribute to new, objective, concept-centric assessment (quantitative vs. qualitative). Precise computational analysis and the classification of the assessment results in terms of Concept States brings a new level of nuance in estimating learner knowledge and in simplifying the knowledge assessment of learning. The structure of the Concept States, classified by the three proposed methods associated with Bloom's Taxonomy, simplify and narrow down the assessment of knowledge domain in one space. The proposed mathematical approaches provide an accurate estimation of the probabilities of knowing or not knowing the concepts in addition to the estimated probability of what the learner is ready to know.

The following section details briefly the publications featuring my research:

1. Chapter 3 has already been published:

    R. Aboalela, J. Khan, "Visualizing Concept Space of course content," IEEE 7th International Conference on Engineering Education (ICEED). pp.160-165. Japan, November 2015. DOI: 10.1109/ICEED.2015.7451512

2. Parts of chapter 4 and chapter 7 have been included in a paper that has already been published by the 2016 IEEE 8th International Conference on Technology for Education (T4E 2016).

    R. Aboalela, J. Khan, "Are we asking the right questions to grade our students in a knowledge-state space analysis?" 2016 IEEE 8th International Conference on Technology for Education (T4E 2016). pp. 144 - 147, Mumbai, December 2016. DOI: 10.1109/T4E.2016.037

3. Parts of chapter 4 and chapter 7 have been included in a paper that has already been accepted by the 3rd International Conference on Soft Computing and Machine Intelligence (ISCMI 2016).

    R. Aboalela, J. Khan, " Model of Learning Assessment to Measure Student Learning: Inferring of Concept State of Cognitive Skill Level in Concept Space," *2016 Third International Conference on Soft Computing and Machine Intelligence (ISCMI)*. Dubai, November 2016.

1.5 Potential Contributions/Benefits

1. Providing visualization of concept-mapped Knowledge Space of a learner

2. Proposing a new, objective, concept-centric assessment (quantitative vs. qualitative)
3. Motivating educators to design new measurement scales using the graph properties (to measure how much is known in terms of the depth, breadth, centrality, etc.)
4. Providing constructive assessment learners can use it as a reference tool for fixing issues
5. Simplifying the difficulties of the knowledge assessment design and learning process through the proposed Assessment Theory
6. Creating a valuable tool for the management of future ACM/IEEE CS curricular revisions, which are expected to have a continued emphasis on Bloom's Taxonomy [1]

1.6 Organization of this Dissertation:

The central focus of this dissertation is the design and evaluation of a learner's knowledge assessment. Furthermore, this dissertation is structured as follows:

Chapter 2 consists of the background research and some related work. The background covers the Bloom's Taxonomy and the Revised Bloom's Taxonomy (RBT). It also mentions related work in the areas combining the Revised Bloom's Taxonomy

---

[1] The Joint Task Force on Computing Curricula Association for Computing Machinery (ACM) IEEE Computer Society. http://www.acm.org/education/CS2013-final-report.pdf

with Knowledge Space, testing and evaluation by using intelligent design, and the Knowledge Assessment Theory (KAT), which provides the probabilities of knowledge states of the assessed individual.

Chapter 3 describes the construction of the Cognitive Level Mapped Concept Graph (CLMCG) as well as the concept-mapped test and evaluation method. It also includes a visualized example of CLMCG of a textbook of the data structure (Ford & Topp, 2002).

Chapter 4 includes the theories used to produce the zones of the Concept States. Chapter 5 describes the estimation of the probability of knowing the concepts and the proposed equations to calculate the probability of knowing the concepts and the learning object in the form of the Concept States. Chapter 6 provides estimation of the Concept States of learners participate in a human subject test

Chapter 7 proposes two experiments; the first of which aims at validating the theories introduced in Chapter 4. The second experiment aims also at validating the theories introduced in Chapter 4, but also describe testing the proposed mathematical equations to estimate the probabilities of knowing the concepts in a human subject test, as introduced in chapters 5 and 6. Chapter 8 contains the conclusion and identifies areas of future research.

CHAPTER 2

Background and Related Work

This chapter introduces the background of Bloom's Taxonomy, the related research work in the field of Knowledge Assessment Theory and testing and evaluation by using intelligent design.

2.1 The Bloom's Taxonomy Background & Revised Bloom's Taxonomy (RBT)

Bloom's Taxonomy (BT) was developed in 1956 under the leadership of educational psychologist Dr. Benjamin Bloom with a view to promote higher forms of thinking in education, such as analyzing and evaluating, rather than just remembering facts (Anderson, et al., 2001). The major idea of the taxonomy is that what educators want learners to know is encompassed in statements of educational objectives arranged in a hierarchy of complexity. The levels are understood to be successive, so that one level must be mastered before the next level can be reached (Anderson, et al., 2001). During the 1990's, Bloom's Taxonomy was revised. Basically, Bloom's six major categories were changed from using nouns to verb forms — for example, the terms "comprehension" and "synthesis" were retitled to "understanding" and "creating." (Anderson, et al., 2001). The main changes to the original BT were that the subcategories of the six major categories were replaced by verbs, and that some subcategories were reorganized. For instance, the word "knowledge" was replaced with the word "remembering" to be appropriate to describe a category of thinking. The revised Bloom's

Taxonomy provides a clear, concise visual representation of the alignment between standards and educational goals, objectives, products, and activities (Krathwohl, 2002). Table 2.1 illustrates the noun-to-verb changes from the original Bloom’s Taxonomy to the revised Anderson’s version.

Table 2.1 The Changes from the Original Bloom’s Nouns to the Version of Anderson’s Verbs

Level #	Bloom's Version Old Version	New Version Revised Bloom’s Taxonomy Anderson’s Version Verb
L6	Evaluation	Create
L5	Synthesis	Evaluate
L4	Analysis	Analyze
L3	Application	Apply
L2	Comprehension	Understand
L1	Knowledge	Recall

2.2 The Knowledge State Assessment Theory (KAT)

In this section, I provide an outline of the development of Knowledge State Assessment Theory, which has taken several years and plenty of research works. The Knowledge State Assessment Theory was employed by Falmagne Cosyn, Doignon, & Thiery (2003). This theory is derived from the Knowledge Space Theory (KST) earlier proposed by Falmagne & Doignon (1999). The theoretical concept of the fringe of a Knowledge State was first used by Dowling, Hockemeyer, & Ludwig (1996), to introduce an adaptive assessment and training program, using the domain of Knowledge States. Using this past research, Flamagne, Doignon and Cosyn (2003) established

Assessment and Learning in Knowledge Spaces (ALEKS)[2]. In 2007, Falmagne and his colleagues established validation and reliability of the theory as a tool for an assessment of scholarly material. Later, Chen (2013), tried to reduce fringe computing time and the number of tested items by presenting pretests to learners. The Knowledge State Assessment Theory was studied by Falmagne (Falmagne J.-C. , Cosyn, Doignon, & Thiery, 2003). The basic two conceptual parts of the theory are Knowledge State and knowledge structure. A Knowledge State is the complete set of problems that an individual can solve on a topic, such as Arithmetic or Elementary Algebra. Knowledge structure refers to an illustrious collection of Knowledge States. The result of the assessment is two lists — "what the learner can do" and "what the learner is ready to learn". "What the learner can do" refers to the most advanced problem in the Knowledge States. "What the learner is ready to learn" refers to the problems that have not been tested, but are more advanced than the problems within the list of the things learners can do. The Knowledge State assessment is a set-theoretical framework deriving from the KST (Doignon & Falmagne, 1999), which proposes mathematical formulae to operationalize knowledge structures in a domain. The most basic assumption of KST is that every knowledge domain can be represented in terms of a set of domain problems (test questions) or items (learning objectives). KST assumes dependencies between the items, in that knowledge of a given item or a subset of items may be a prerequisite for

[2] ALEKS® is a registered trademark of ALEKS Corporation.©2015 McGraw-Hill Global Education Holdings, LLC. All rights reserved.

knowledge of another, more difficult or complex item. These prerequisite relationships are realized through surmise relations, which create a quasi-order between different items. However, they do not illustrate the semantics of the relationships. In our research, we identify a pedagogical relation between the concepts in one space, and RBT labels these relationships semantically. As mentioned by Stahl (2011), “One advantage of the surmise relations is that they reduce the quantity of all possible solution patterns to a more manageable amount of Knowledge States. Each of these Knowledge States represents the subset of items an individual can solve. The collection of all Knowledge States captures the organization of the domain, and is referred to as a knowledge structure”.

In their model, Falmagne, Cosyn, Doignon, & Thiery (2003) use the example of states defined as whether a learner “knows” or “does not know”. In our study, the researchers modeled the states in a different way: (know, does not know), derived (indirectly known), derived unknown (indirectly inferred is not known), potential (ready to know), and potential unknown (not ready to know). More accuracy was added to the states in the assessment of the Knowledge Space by identifying cognitive skills relationship between the concepts. In KAT, it is considered that in any Knowledge State there is a fringe, which is a set of items by which the Knowledge State differs from its neighbors. The fringe of any Knowledge State involves two areas, namely the outer fringe and inner fringe. The two terms describing the fringe in KST were introduced by Falmagne & Doignon (1988). Later, Doignon and Falmagne (1999) examined the concept of a neighborhood, which means if the symmetric set difference between two

Knowledge States is at most one, they are considered neighborhood states. From this point the concept fringe of a Knowledge State is defined as the set of items by which the Knowledge State differs from its neighbors. In other words, fringe differentiates the Knowledge States of neighborhood items. The outer fringe of any Knowledge State contains the set of more advanced problems that were not included in the basic Knowledge State. The problems within the outer fringe refer to the Knowledge State, "the learner is ready to learn." The problems within the inner fringe refer to the Knowledge State, "the learner knows", which means that he/she gets high points on these items. Dowling, Hockemeyer, & Ludwig (1996) employed the concept of the fringe of Knowledge State for adaptive assessment. The researchers have used the concept fringe to differentiate between the Knowledge States of neighboring items (Dowling, Hockemeyer, & Ludwig, 1996). They tried in their study to further reduce the computing time and memory of the Knowledge Space based procedure as compared to the studies by Falmagne & Doignon (1988) and Hockemeyer (1993). The number of tested items used to complete the assessment was the lowest in the work of Dowling, Hockemeyer, & Ludwig (1996). The main idea of their work was to exploit the theoretical concept of the neighborhood of a Knowledge State in the adaptive training.

The work of Falmagne, Doignon and Cosyn (2003) is reality, and accomplished by establishing an Assessment and Learning in Knowledge Spaces (ALEKS)[3], which is a

[3] http://www.aleks.com/about_aleks

web based, artificially intelligent assessment and learning system. As explained in (ALEKS)[4], ALEKS uses adaptive questioning with a purpose to quickly and accurately find exactly what a learner knows and does not know in a course. ALEKS could guide the learners on the topics they are most ready to learn. The assessment then provides lessons for the concepts they are ready to learn. The system introduces to the learner an adaptive assessment to compute his current Knowledge State based on Bayesian likelihood logic over the learning space. The resulting values imply the probability of what questions the learner can answer even though they may not have been tested yet. In 2007, Falmagne and his colleagues established the validity and reliability of the Theory as a form of assessment of scholarly topics (Falmagne J. C., Cosyn, Doble, & Uzun, 2007). The most current research employing Knowledge Space for adaptive assessment is the work of Chen (2013), who tried to find a dynamic procedure for computing the fringes of a Knowledge State according to the prerequisite relationships between knowledge nodes (competency). Adaptive assessment can be defined as a form of assessment, where the items on the test are tailored based on the individual's performance (answer) on previous items. The researcher does not specify the type of relationship between the competencies but assumes the prerequisite relations are known. The effort in Chen's (2013) study was to present a procedure to reduce the fringe computing time and number of tested items by presenting the pretest to the learner. Chen proposed the

[4] ALEKS® is a registered trademark of ALEKS Corporation.©2015 McGraw-Hill Global Education Holdings, LLC. All rights reserved

procedure based on the two fringes of the Knowledge State. The two fringes were earlier used in adaptive assessment by Dowling, Hockemeyer, & Ludwig (1996). As Chen explained, the pretest is the first item introduced to the learner in the test. The result of the pretest directs the assessment to be more advanced to test the outer fringe, or less advanced to test along the inner fringe. The result of the assessment is supposed to show whether the learner's Knowledge State is confirmed mastered or unconfirmed mastered. Chen used a competency to represent a type of basic mathematical problem (learning object), and used sub-competencies to represent the basic problems in the same way as Auzende, Giroire, & Le Calvez (2009).

2.2.1 Discussion

The main problem in assessing the knowledge in one domain is to find the Knowledge State of the assessed individual in the minimum time and question him/her. The basic information in Knowledge Space Theory (KST) (Doignon & Falmagne, 1999, pp. 247-252) provides the Questioning Rule and Marking Rule to reduce the number of questions needed to be asked to discover the Knowledge State of the individual. Despite the sophistication of the Rules, they are appropriate for use in the Knowledge Space due to their clear illustration of the relationships and strong dependencies between the concepts in domains such as algebra, math and chemistry. In these learning spaces, the dependencies are clarified by the basics needed to solve the math equation. Accordingly, the prerequisite relation between the competencies is known by the sequence of the complexity of the related questions. The sequence thus derived from the simple question includes everything from the simpler and elementary knowledge, to the more complex and professional knowledge. In KAT, Falmagne J.-C., Cosyn, Doignon, & Thiery (2003) by practice assume prerequisite relation between the topics, which occur together in one subject to make a feasible Knowledge State. As in Pre-Algebra, their introduced example, the topic of long division always includes whole number subtraction. In knowledge assessment such as KAT, the dependency relationship between the concepts does not illustrate the semantics of these relationships, but assumes a pedagogical relationship between the concepts in one space. This deficiency in KAT methodology restricts or eliminates the assessment under a complex knowledge structure. Moreover, in disciplines such as computer science and software engineering, the relationship between the concepts should be identified clearly.

Our proposal to counter these problems is contained in Chapters 3 and 4, where we provide a model of CLMCG and identifying the Concept States organized by the skill levels of the concepts.

Chen (2013) introduced rules to reduce the computing time of the fringes of a Knowledge State, based upon the prerequisite relations between knowledge nodes. She used the links connecting the competencies rather than the links between the Knowledge States.

2.2.2 Differentiation

One of the improvements made by Chen to previous work in the area is in the consideration of the prerequisite relation as a direct link between the independent competencies. The relationships in Chen's work imply the relationship between the competencies, whereas, in KST (Doignon & Falmagne, 1999), the link was understood to be existing between the Knowledge States. The biggest addition is that Chen presents each competency as a node and identifies a direct link between these competencies with a purpose to reduce the computing time of the fringe. Chen does not identify the knowledge structure, nor the independent competencies. Similarly, Chen also does not identify the relationships between the competencies, but instead assumes that these relationships are known within the assessment domain. However, Chen does identify the rule to discover the fringe and to select the question from the closest node to the fringe. While Chen doesn't validate the relationship between the prerequisite competencies in the learning path.

This study is differentiated from Chen's by identifying clear pedagogical relationships between all the concepts in the domain. Chen didn't identify the relationships between the competencies, but instead these were assumed known in domains like in the math domain. The prerequisite relationship in domains like math may be classified from the less advanced competencies to master the more advanced competency. In domains with less dependent competencies, such as the domain of Computer Science and applied science, identifying the prerequisite relationship is important to identify the connection between the competencies. In Chen's work, the fringe is computed from the discovered Knowledge State of the first pretest, and then the test consists of all the concepts the path controls through the answered questions. We contribute by classifying the Knowledge State in the form of Concept States, which simplify the knowledge structure and reduce the number of tested concepts. The tested concepts are classified as the zones of Verified Skills which contains the set of the advanced learning objects, and would direct the path of the assessment in minimum time. The learning object concepts in the set VS replaces the purpose of the inner fringe. DS, which replaces the purpose of the outer fringe, may be estimated from the VKS and VNS.

2.2.3 Improvement

Analysis methods are proposed to discover the Concept States of the assessment. The most differentiated aspect of this study is that the Concept States are identified with a taxonomy link between all the concepts in the assessment domain. The link is a verb identifying the certain skill needed to be learned in the prerequisite concept and to be able to attain the target concept at certain skill. The knowledge structure in KAT is improved by adding the Concept States, which classify the concepts based on the answer to the tested concepts at the specified skills.

2.3 Combining the Revised Bloom's Taxonomy with Knowledge Space

There is some research associating the knowledge structure in Learning Space Theory with skills and competences (Doignon J.-P. , 1994), (Heller, Mayer, Hockemeyer, & Albert, 2005), (Heller, Steiner, Hockemeyer, & Albert, 2006), (Marte, Steiner, Heller, & Albert, 2008), and (Reimann, Kickmeier-Rust, & Albert, 2013). The research trying to associate the KST with Bloom's Taxonomy is the work of Marte, Steiner, Heller & Albert (2008), where they used the ideas of Competence-based Knowledge Space Theory (CbKST) following (Heller, Steiner, Hockemeyer, & Albert, 2006). They propose a skill that can be characterized as a pair consisting of a concept and an activity. As an example of such a pair, they give "*Apply* the *Pythagorean Theorem*". In our work, we concentrate on the concepts as they appear in the text in either phrase form or single word form and identify the link between the concepts as the skill required to learn the concept at a certain skill level, which identifies the prerequisite relation between the concepts. The researchers Marte, Steiner, Heller, & Albert (2008) presented a knowledge representation model that can incorporate the activity-oriented understanding of teaching and learning. They associate skills with the problems and learning objects by presenting two maps: skill functions and problem functions. In skill-function mapping, they mapped a collection of the subset of skills to the assessment problems. Each subset's competencies consist of skills sufficient to solve the problem. The collection of skills a person has available is called the competence state of the individual. If the Knowledge State of the learner identifies low-level learning objectives, this may indicate the skills to be learned next. The objectives identify the knowledge and competence state of the learner. Each problem could be mapped to many competencies since it could be solved in more than one way. In a problem-mapped function, the researchers mapped each subset of skills to the set of problems. They then characterized the skills as a pair consisting of a concept and an activity, as the given example "*Apply the Pythagorean Theorem.*" The pair illustrates the prerequisite relation between the basic concepts by using the concept map tools. In their work, the skills are meant to provide a fine-grained, low-level description of the learner's capabilities. Figure 2.6.a and 2.6.b illustrate the two mappings. As may be seen, they separate the skills and learning objects. Subsequently, the mappings connect the skills based on Bloom's Hierarchy, and the learning objects based on the prerequisite attributes. Figure 2.6.c shows the final graph that combines the skills and learning objects in one node.

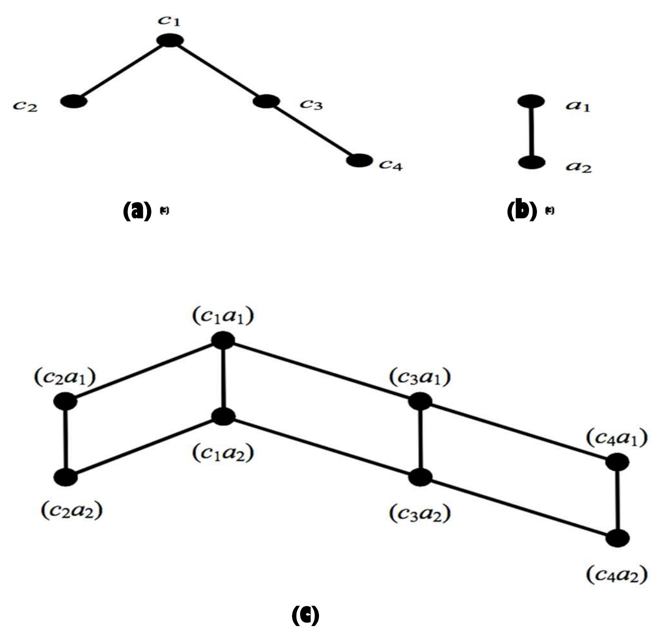


Figure 2.1 The Mapping According to the Competence-based Knowledge Space Theory (CbKST)

2.3.1 Discussion

Our research intersects the following points:

1. Combines Revised Bloom's Taxonomy with Knowledge Space
2. Uses conceptual ontology to connect the concepts
3. Associates the concepts with the verbs of Bloom's Taxonomy

2.3.2 Differentiation

We differentiate from past research by concentrating on the concepts as they occur in the text, not as associated with the skill verbs. The main concept of past research has been the skill of the Bloom's Taxonomy and the subset as competence. The researchers did not identify skills as the links between concepts or between learning objects. Identifying the links with the verb of the skill, that connects the concepts in the context, increases the precision of the assessment and controls the direction of assessment. In this study, the verbs of skills (skills verb) are used to link the concepts, which means the verb is identified by the link, and not associated with the evaluated concepts.

2.3.3 Improvement

This work complements the work of Marte, Steiner, Heller, & Albert (2008), which combines RBT with Knowledge Space Theory (KST) (Doignon & Falmagne, 1999), by adding the skill levels of RBT to connect the concepts in the assessment with the concepts in the assessed domain, rather than considering the concepts as skills and recognizing the ontology relation as is done in (Marte, Steiner, Heller, & Albert, 2008). We have proposed analysis methods to discover the Concept States of the assessment. This makes a suitable complement as well as an addition to the existing work combining Knowledge Space with RBT.

2.4 Testing and Evaluation by Using Intelligent Design

The research works of (Khan & Hardas, 2013), (Khan, Ma, & Hardas, 2006) (Khan, Hardas, & Ma, 2005). In (2005) the authors studied the problem of difficulty in evaluation by using Semantic Network Ontology-based Intelligent Courseware Sharing. Their research was aimed at automating the process of intelligent design by using test-ware and providing a qualitative assessment of questions. The researchers provided some synthetic parameters for the evaluation of questions in Concept Space. They observed that the difficulty of a question is often a function of the concepts that it tests. They claimed that the concept knowledge could be represented in the form of linked concepts in semantic nets, where the links represent the relationships between the concepts. If this directed graph is known, the complexity of a question can be computed by synthetic means. Khan et al., (2005) introduced the Topic Dependency Graph (TDG), which is a projection of a semantic net for making assessments on a course. The TDG is further associated with a weight system. The self-weight represents the relative semantic importance of the root topic with respect to all other prerequisites. The prerequisite weights represent relative semantic importance among the prerequisite topics. The TDG gives the layout of the course in the Concept Space, while also specifying the course organization, involved concepts, and the relations between the concepts. Using this foundation, Khan et al. (2005) introduced test design and evaluation. Their evaluation model's basic unit is problematic, and it is connected to a concept set from the ontological representation of the course. This is known as a question in concept mapping. The ontological link between the semantic concept pair is classified as "AND" or "OR."

The concepts which are essential requirements are linked by the “AND” relation, while those which are not imperatively required are shown by the “OR” relationship. Their effort was to quantify the amount, which is tested by the problem using the concepts individually and with respect to the ontological root.

2.4.1 Discussion

Our research intersects in the following points:

1. Provides a platform for making a transfer from quantity assessment to quality assessment on the concept-centric dimension
2. Facilitates the assessment process by representing the assessed domain in the form of linked concepts
3. Connects the test question to the existing assessed domain
4. Provides the evaluation scales and test design

2.4.2 Differentiate

The paradigm graph CLMCG is proposed to simplify the complexity of the test-question process. The weight of the prerequisite concepts is not added in, but instead the RBT levels are labeled based upon the conceptual ontology. Thus, we designed a paradigm CLMCG that is simpler and more intelligent, to provide a flexible graph design for any course study to be assessed. We claim the possibility of labelling any course concepts’ ontology and present a simple model to simulate CLMCG.

2.4.3 Improvement

We complemented the work of Khan et al. (2005) by adding RBT levels. Labeling the course Conceptual Ontology with RBT will control the concepts' dependency. The ontological link between the semantic concept pair is classified as "AND" or "OR," and will be specified as to the level in the concept that is most important. We linked the conceptual ontology according to the real relation ontology component rather than the prerequisite of the ontological links. There are some research works trying to take the benefit of the connected domain with RBT such as that of Nafa, Khan, Othman, & Babour (2016 a, 2016 b) and Nafa & Khan (2015). They identify methods by using the verbs in the text to discover and utilize the cognitive relation between the concepts. In our proposal works, we identify the cognitive relations between the concepts by using the skill levels of RBT. We take an advantage of the cognitive relation to identify the zones of the proposed concept states. The zones of the concept states simplify the estimation of knowing the concepts. Also, using the Bayes' Theorem in the mathematical computation increased the accuracy of the estimated concept states of the learner. We add numerical methods to identify the probabilities of knowing and not knowing the concepts of the learners in the evaluated domain.

CHAPTER 3

The Cognitive Level Mapped Concept Graph (CLMCG)

In this chapter, we discuss the model of the CLMCG and the visualized implementation of it from a specific textbook. This model is a part of the Theory of Cognitive Skill in Concept Space (TCS^2). TCS^2 theory is an idea that presents a solution to ascertain the pedagogical relationship between existing concepts in one domain to help objectively (algorithmically) understand and assess the learning states and skill levels of a learner, with respect to the conceptual contents, concepts and relationships that define a specific knowledge domain called Concept Space. The theory is composed of four components: the representation of the assessment domain in the aspect of mapped concepts space called Cognitive Level Mapped Concept Graph (CLMCG), the concept-mapped testing and evaluation method, the Concept States in the Theory, and the TCS^2 assessment analytics, which identify the process to estimate the sets of the Concept States at cognitive skills levels. In this chapter, we discuss the implementation of CLMCG, and the concept-mapped testing and evaluation method. Part of this chapter's content is also published on (Aboalela & Khan, 2015). Also, the visualized implementation is realized on the website[5].

---

[5] http://rania.medianet.cs.kent.edu:8080/Project

3.1 The Cognitive Level Mapped Concept Graph (CLMCG)

The Cognitive Level Mapped Concept Graph (CLMCG) is a concept-mapped knowledge domain that organizes the concepts of a knowledge domain (such as computer operating system, data structure, etc.) into a matrix structure with three relational dimensions: the syllabus dimension, the ontology dimension and the cognitive skill dimension. The syllabus dimension retains the occurrence of the concepts in a formal textbook such that it retains the chapter, section, sub-section, etc. The ontological dimension links the concepts in terms of class, part, and instance of relationships. The cognitive skill dimension captures the relationship between concepts in terms of support needed from one concept to another concept to attain a level of skill in said concepts. The cognitive skill dimension RBT links to the ontological dimension. In other words, the cognitive skill dimension is the mapped-knowledge domain with RBT. We have mapped the syllabus dimension to the ontological dimension to obtain the mapped knowledge domain. Then, we added the RBT analyzing over the mapped knowledge domain to obtain cognitive dimension and accomplish the CLMCG.

3.1.1 The Syllabus Dimension [Area Knowledge Space]

An area Knowledge Space contains all the elementary concepts covered in the knowledge domain. A textbook or a syllabus provides organization for the concepts in a tree hierarchy. There can be multiple textbooks, which can organize hierarchy in different ways. We call each of these trees a “Concept Organization Schema (COS)”. For the following discussion, we will only assume the concept organization scheme given by one book and call it the total syllabus of the AKS.

The objective of COS is to capture the presentation and organization of the AKS provided by the author of the textbook or syllabus. Thus the root of the AKS has the name of the course. The leaves of the AKS are the fine grain concepts. The internal nodes represent the chapters, sections, sub-sections, paragraphs and sentences in which the concept appears. Thus, the AKS tree has only one type of link, i.e., "occurs-in." The nodes, however, have types, such as concept, section, paragraph, chapter, etc. and have a concept title. It should be noted that a concept could appear in multiple places. We used separate node instances to keep track of each occurrence of the concept, and kept the general tree structure intact. The Location Identifier (LID) of each concept occurring inside a reference text was used to keep track of the concept track. One concept may occur in multiple places and thus may have multiple Location Identifiers. This is a hierarchical ID identifying everything from the root to the concept of the following format: [Book ID. Chapter#. Section order #.Subsection order #.Paragraph order #.Sentence order#. Concept order#]. However, as we will explain later, ontological dimension is used to derive such equivalence.

3.1.2 The Ontological Dimension

The ontological dimension captures the ontological relationship between concepts. We identified the ontological relationship between the concepts based on

WordNet[6], and the interrelationships existing for a Concept Space. We represented the links of four relationships between the concepts as: Sub-Part (SP), Sub-Class (SC), Synonym (Synm), and Instance (IN). The relationships are represented in the graph as directed dashed links from the source to the target. e.g. if A is, Sub-Part, Sub-Class, Instance of, or Synonym of concept B then concept A is represented as the source node and the concept B is represented as the target node. Figure 3.1 illustrates the ontological relationship. L_k refers to the ontological relationship type.

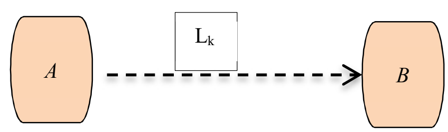


Figure 3.1 The Ontology Map

3.1.3 The Cognitive Dimension Connection

The cognitive dimension connection is the Bloom's Taxonomy analysis of the ontological dimension. The cognitive skill dimension captures the relationship between concepts in terms of prerequisite concept needed to be attained in order to know the target concept at a certain skill level. The Cognitive Skill Level refers to levels needed to acquire the concept at the cognitive levels of understanding, Applying, analysing, or creating after knowing the prerequisite concepts. We eliminated the Cognitive Skill Level

[6] We used the WordNet relation sets. But the concepts in our CLMCG is not identically related with the actual WordNet.

on the graph by using the five higher verbs of Bloom's Taxonomy. To map the computer science knowledge domain with RBT, we added labels to RBT verbs. We also added new verbs, which are specifically used to describe computer science knowledge acquisition. Thus, the cognitive dimension could have an unlimited number of cognitive skills. The cognitive level relation is represented in the graph by direct edge from the source concept to the target concept. The source concept is the prerequisite concept and the target is the dependent concept. Figure 3.2 shows an example of the cognitive level map. The direct solid edge from concept A to concept B means that to understand concept B, the cognitive level of concept A must be known. Lk refers to cognitive level (verb). However, there are some researchers who try to combine in the learning assessment, the Knowledge Space with activity and Bloom's Taxonomy (Marte, Steiner, Heller, & Albert, 2008), (Albert & Held, 1999). Our most differentiated aspect is that we identified the Concept States associated with Taxonomy links which connect all the concepts in the assessment domain. The Taxonomy link is a verb that identifies the skill that could be learned in a the target concept. In RBT, the same verb can be used to indicate different Cognitive Skill Levels. We addressed this problem by assigning a label number to the taxonomy link, which indicates the main verb and the subcategory. The format of the label is as follows: (RBT verb number. the subcategory verb number). For example, the skill link Lk = 4.18, where 4 indicates level 4 in RBT, and 18 indicates the subcategory verb from computer science, which is the verb "scan" in our implementation of the list of verbs.

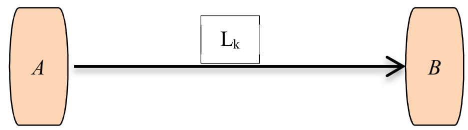


Figure 3.2 The Cognitive Level Map

3.1.4 Visualized Example of CLMCG

Currently, in our CLMCG, we represent the concept occurring in Data Structures with C++ using the STL (2nd Edition) book (Ford & Topp, 2001). We implemented the presented visualizing graph from this book, and found some statistics based on the three dimensions. Table 3.1 & Table 3.2 respectively show the overview statistical and distributed degree of the three dimensions of CLMCG. Figures 3.3 (a) (b) and (c) respectively show snapshots of the AKS, and the ontological and cognitive view. Figure 3.3 (d) shows a zoomed-in view of the cognitive relationship of one concept, “Selection sort algorithm.” The real visualized example is implemented in the website rania.medianet.cs.kent.edu:8080/Project/#.

Table 3.1 The Statistic of the Cognitive Level Mapped Concept Graph

A. The Syllabus Dimension	
Number of hierarchal links based on syllabus organization.	17510
Number of nodes with frequencies	4284
Number of total Concepts without frequencies	400
B. The Ontological Dimension	
Number of entire ontology relations	998
Number of Synonym, Synm Ontology Relation	400
Number of Sub Class, SC Ontology links	104
Number of Sup part, SP Ontology links	486
Number of Instance, IN Ontology links	8
C. The Cognitive Dimension	
Number of entire cognitive relations	1373
Number of L2 cognitive relations	88
Number of L3 cognitive relations	449
Number of L4 cognitive relations	350
Number of L5 cognitive relations	174
Number of L6 cognitive relations	312

Table 3.2 The Statistic of the Graph Diameter & Cluster Coefficient of the Cognitive Level Mapped Concept Graph

A. The Syllabus Dimension	
Graph Diameter	14
Clustering Coefficient	0.668
B. The Ontological Dimension	
Graph Diameter	7
Clustering Coefficient	0.368
C. The Cognitive Dimension	
Graph Diameter	11
Clustering Coefficient	0.15

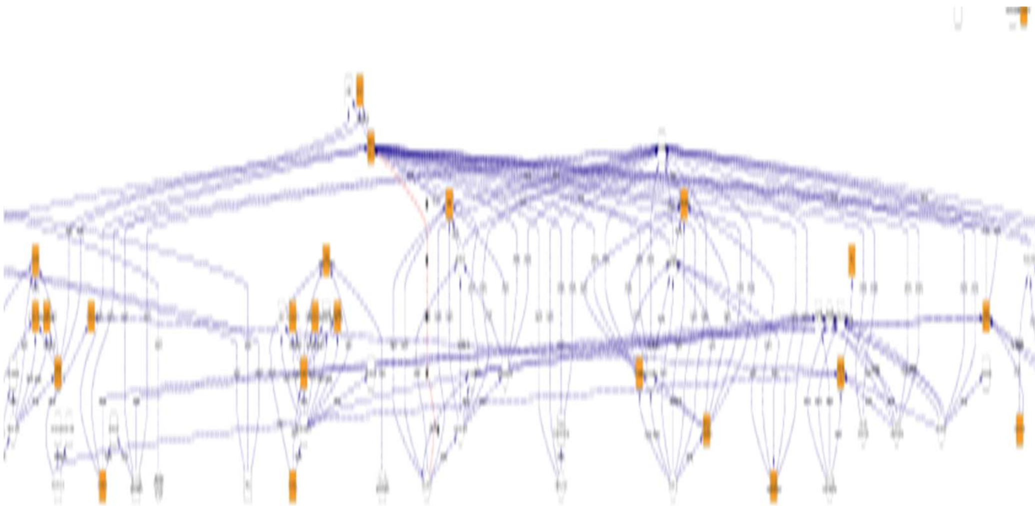


Figure 3.3 (a) Zoom out Snapshot of the Syllabus Dimension (The Hierarchal View)[8]

---

[7] This visualization is implemented by using open source program called Cytoscape. http://www.cytoscape.org

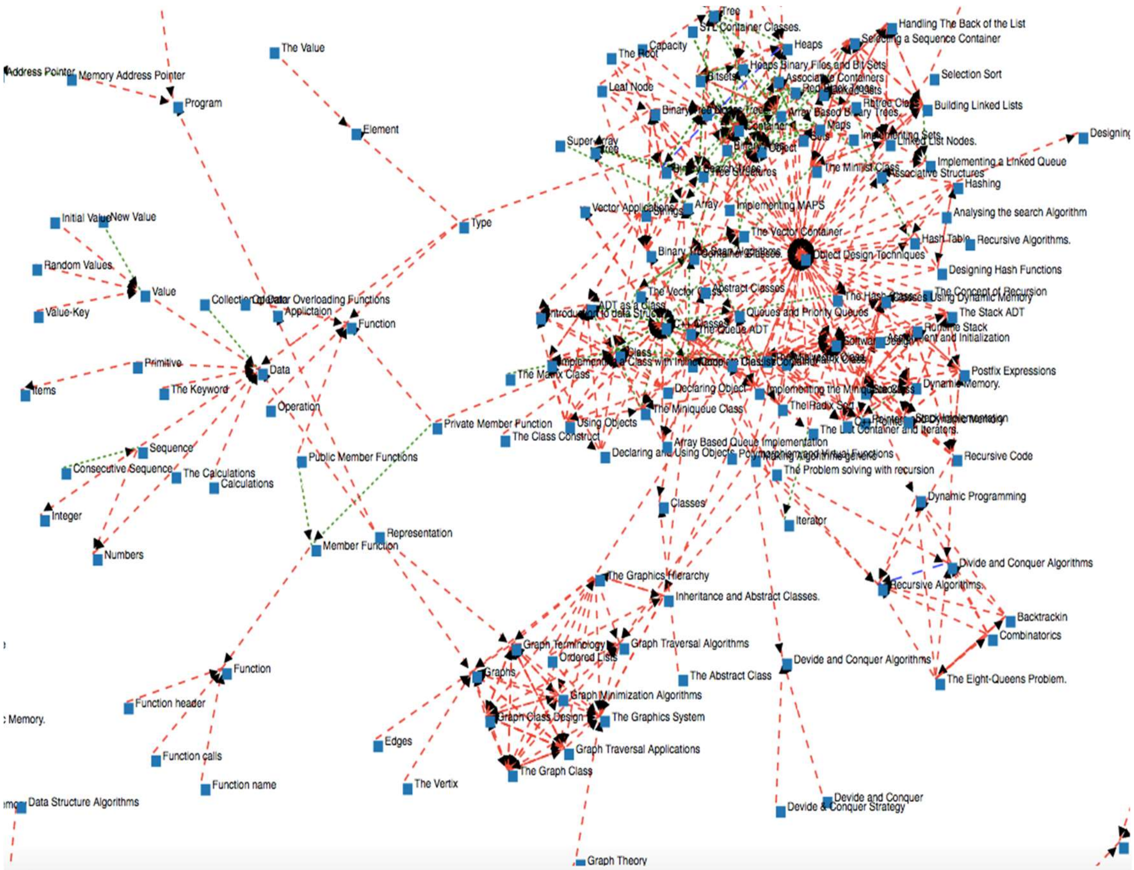


Figure 3.3 (b) Zoom out Snapshot of Ontology Concept Map[9]

---

[8] This visualization is implemented by using an open source program called cytoscape. http://www.cytoscape.org

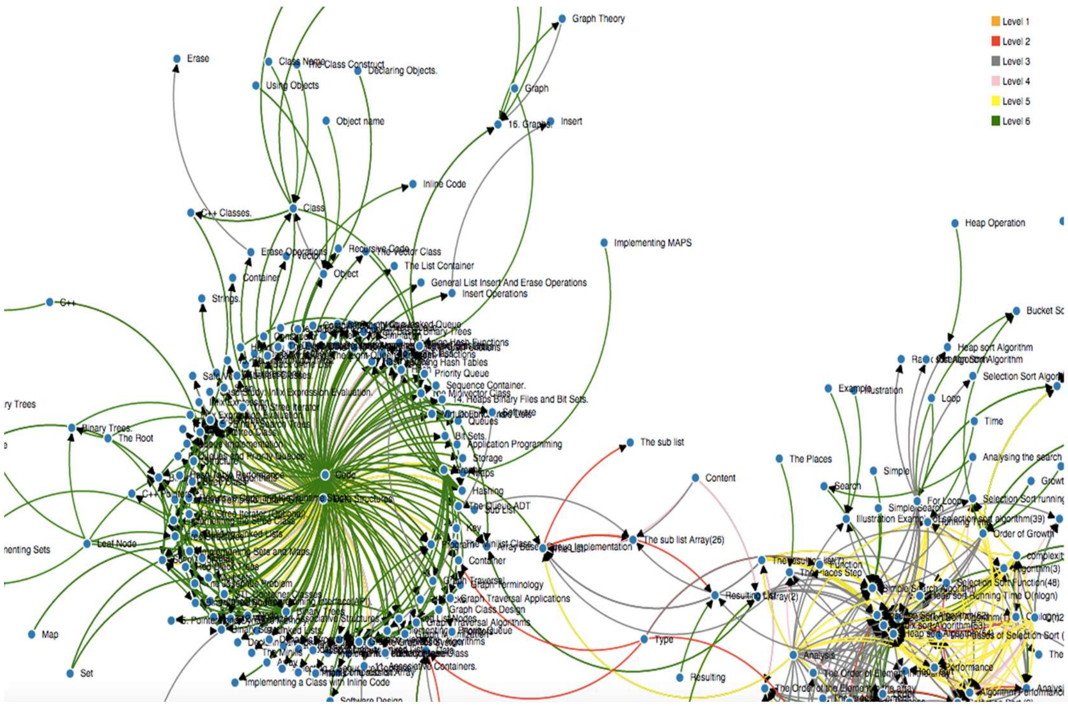


Figure 3.3 (c) Zoom out Snapshot of Bloom's Links of Computer Science Concepts[10]

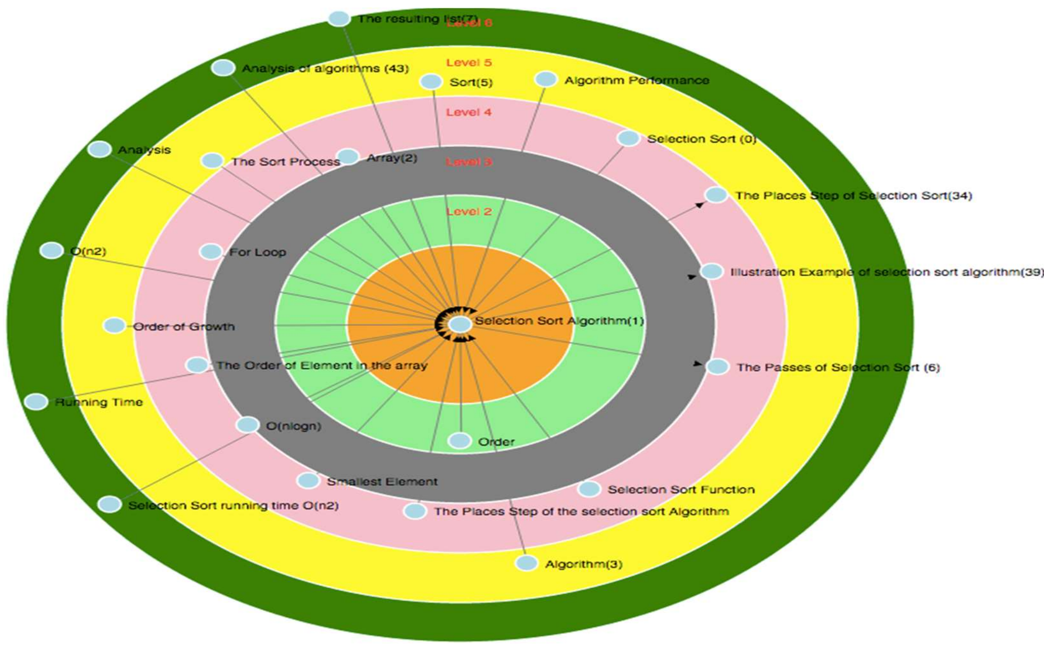


Figure 3.3 (d) Zoom in Snapshot of the Concept "Selection Sort Algorithm" Cognitive Level Mapped Concept Graph

---

[9] This visualization is implemented by using D3, the Javascript library for visualization tools in web browser

3.2 Concept-Mapped Test and Evaluation Method

In order to measure the learner's learning, we set up concept-based testing and evaluation methods. In conventional testing, a learner is given a set of questions. A grader evaluates the answers and assigns a quantitative score to the learner. We slightly modified this evaluation method, wherein the grader is asked to evaluate whether there is evidence in the answer that the learner has succeeded or failed to attain a certain Cognitive Skill Level, instead of the usual numerical score. We called it "Concept-Mapped Evaluation". Sometimes the questions can also be specially designed to directly measure the skill level of the certain concepts. We called this "Direct Concept-Mapped Testing". The direct questions address the identical skill level targetted by the instructor, and directly specify the level of the concept. Therefore, two types of questions are offered: (1) open questions and (2) direct questions. Open questions are questions without restrictions. An open question is a conventional question prepared by the instructor and could implicitly test the skill level. The direct questions are specially-designed questions which ask exactly about a certain skill of the tested concept.

CHAPTER 4

The Zones of the Concept States

This chapter explains the zones of three basic sets of Concept States: Verified Skills (VS), Derived Skills (DS), and Potential Skills (PS). After the assessed individual completes the assessment his result will be in the form of six Concept States: Verified Known Skills (VKS), Derived Known Skills (DKS), Potential Known Skill (PKS), Verified Not knowing Skills (VNS), Derived Not Knowing Skills (DNS), and Potential Not Knowing Skills (PNS). These six sets are derived from the three basic sets. We made a complement to the Knowledge States of Falmagne, Cosyn, Doignon, & Thiery (2003) and the work of Marte, Steiner, Heller, & Albert (2008); Heller, Mayer, Hockemeyer, & Albert (2005); and Heller, Steiner, Hockemeyer, & Albert (2006), which combines RBT with Knowledge Space Theory (Doignon & Falmagne, 1999) by adding the skill levels of RBT to connect the concepts in the assessment with the concepts in the assessed domain, rather than to consider the concepts as skills and activities, and recognize the ontological relationship as is done in the work of Marte, Steiner, Heller, & Albert (2008). On the other hand, KAT concerns sequential knowledge and the growth of knowledge (i.e., to know concept A, the individual must know the prerequisite concept set), and lacks the determination of the level of skills. We proposed making the link between any two concepts as a verb of the skill which needs to be learned; for example, to Apply a concept A, the individual must know the prerequisite concept set for the appropriate skill levels. Here, “Apply” is in skill level 3 of RBT. We proposed analysis methods to discover the Concept States of the assessment. This makes a sufficient complement to the works combining Knowledge Space with RBT such as Marte, Steiner, Heller, & Albert (2008); Heller, Mayer, Hockemeyer, & Albert (2005); and Heller, Steiner, Hockemeyer, & Albert (2006). The elementary Theory of the Concept States is formed from, for example, a simple link from node A to node B, which means: To [Lk] B one must know A, where "Lk" is a link containing a verb of the skill that connects between two concepts. For example, to Apply "Sort", the learner must know the "Sorting Algorithm." A is the concept "Sorting Algorithm" (from an algorithm course in computer science). B is the concept "Sort". Lk = 3 = "Apply", which is a verb of level 3. In the same way, if we want to ask about analyzing the sorting algorithms, the learner must know running time. Thus, node A is the concept "Running Time", and node B will be the concept "Sorting Algorithm" and the skill level Lk = 4, which is a verb in RBT. Based on this Theory, we propose three basic sets of Concept States: Verified Skills (VS), Derived Skills (DS), and Potential Skills (PS). Note that there are two different definitions for Derived Skills (DS) depending on the level of the state. When a correct answer is given to the question asked about a concept at certain skill level, then the tested concept will be added to the set of the appropriate known state, where the known Knowledge States is one of the three proposed known states: Verified Known Skills, Derived Known Skills, or Potential Known Skills. If the incorrect answer is given to the question asked about a concept at a certain skill level, then the tested concept will be added to the appropriate not known Concept State. The not knowing state is one of the three proposed not knowing states, which are: Verified Not known Skills (VNS), or Derived Not known Skill (DNS), or Potential Not known Skill (PNS).

4.1 Zones of Verified Skills, VS (k)

Verified Skills (VS) are defined as where there is direct evidence that an individual knows or does not know a concept C_x at a Cognitive skill level k. If the evidence is present, that concept is considered to be a part of verified set VS(k). To illustrate the VS, let us consider that we are gathering evidence by a question Q, which can ascertain that a learner knows or doesn't know a specific concept C_x. If a question Q asked about the concept C_x at level k, and the answer of assessed individual is a correct answer, then the concept C_x will be added to the set of Verified Known Skills (VKS). If the answer is incorrect, then the concept would be added to the set of Verified Not Knowing concepts (VNS). Thus, VKS satisfying the condition that: If $(Q_i,C_x)^{Lk}$ & C_x is a correct answer, then $C_x \in$ VKS(k) $\forall$ $C_x \in$ completely correct answer concepts. $Q_i \in$ test questions, $C_x \in$ tested concepts, Lk $\in$ Bloom's Link of level k, VKS(k) $\in$ Verified Skills at level K and $(Q_i,C_x)^{Lk}$ is the 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 4.1 and Figure 4.2 show the Verified Known Skill link and the Verified Not Known Skill link respectively.

For example, a question Q could be: "sort the elements in the array by using selection sort algorithm". The set of verified skills will be the concepts $\{C_1^3, C_2^2, C_3^2\}$. C_1^3 is the concept "Selection Sort Algorithm" at skill level 3 of subcategory verb "Sort"; C_2^2 is concept "Array" at level 2; and C_3^2 is the concept "Element" at level 2. Therefore, the link from the question Q_1 to concept C_1^3 is Lk = 3.20, where 3 indicates level 3 in RBT, and 20 indicates the subcategory verb from Computer Science, which is “Sort” in our implementation of the list of verbs. Similarly, links will be assigned with the appropriate label numbers for the concepts “Array” and “Elements”.

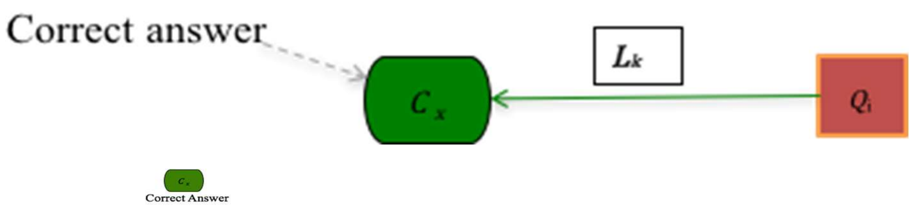


Figure 4.1 Verified Known Skill (VKS) Relation

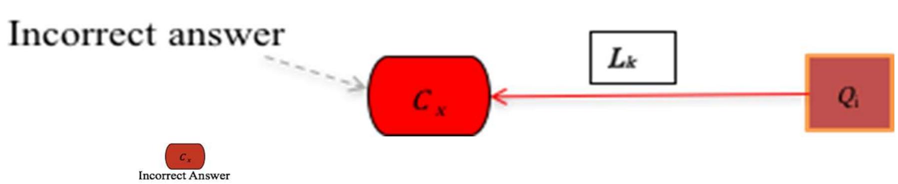


Figure 4.2 Verified Not Known Skill (VNS) Relation

4.2 Zones of Derived Skill at Level 2, DS (k=2)

The DS is defined as a set of concepts in the prerequisite set at certain skill level of the tested concepts but they have never been directly tested. For example, Derived Skill (DS) at skill level 2 is defined as the existence of indirect evidence that the concept C_i is understood or not understood by the learner. It will belong to DS (K=2). The condition of the relation is expressed as the following:

If C_i is not in a verified set, there exists two links such that $(Q_i,C_x)^{Lk}$, $(C_i,C_x)^{Lm}$ & $C_x \in VS(k)$, then it is in VS, and m= 2 and $k \geq m$, then C_i is 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 $(Q_i,C_x)^{Lk}$, $(Q_i,C_x)^{Lm}$ means there is an existing link between the Question Q_i and the Concept C_x at levels k and m, respectively. For example, to Apply the concept C_1, which could be "Selection Sort Algorithm", a learner must understand the concept such as C_2 which could be "order." Thus, the link will be from the concept C_2 to the concept C_1 and the skill link Lm = 2, which indicates the understanding level in RBT. In other words, if the concept C_1 in the set of VS and the concept C_2 is a prerequisite to the concept C_1, then the knowing of the concept C_2 will follow the evidence of knowing the concept C_1. For example, if the concept C_1 = "Selection Sort Algorithm" is an element in VKS, then prerequisite concept C_2 = "Order" will be added to the Derived Known Skill at level 2 DKS(2). If the concept C_1 is an element in VNS, then the concept C_2 will be added to the set DNS(2). Figure 4.3 and Figure 4.4 illustrate the relation in DKS and DNS respectively.

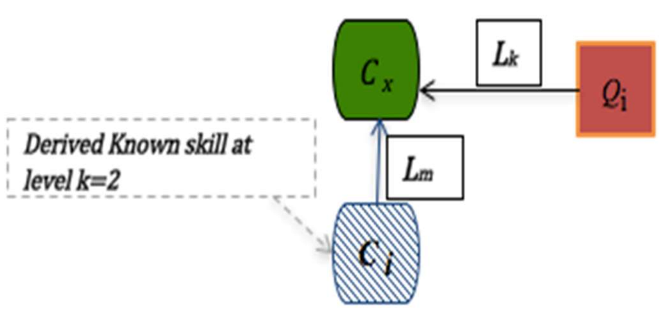


Figure 4.3 Derived Known Skills (DKS) Relation at Level 2

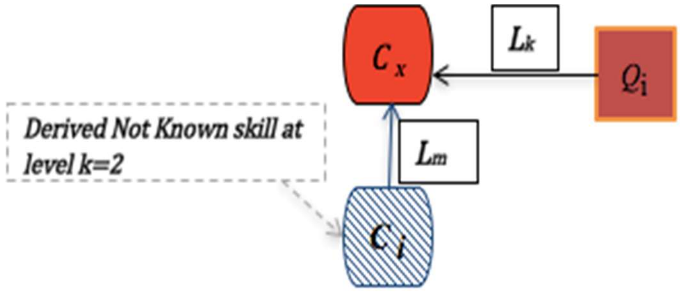


Figure 4.4 Derived Not Known Skills (DNS) Relation at Level 2

4.3 Zones of the Support Set (SS)

To distinguish the higher-level relations, the classification of the set into Support Set and Supported Set must be identified. The support set means the prerequisite set of the supported set.

Let C_A be a node. Let C_B be another node from where there is a level k link to A. Then we call C_B level k the support node of C_A. That means C_B is the prerequisite set of C_A concept at level k.

Let S (C_A, k) be the set of all such C_B nodes in the complete concept graph G. The S (C_A, k) is the level k Support Set for C_A. i.e. all concepts in this set must be learned to have a level k skill in A. Figure 4.5 illustrates the Support Set & Support Node, which is any node in the Support Set.

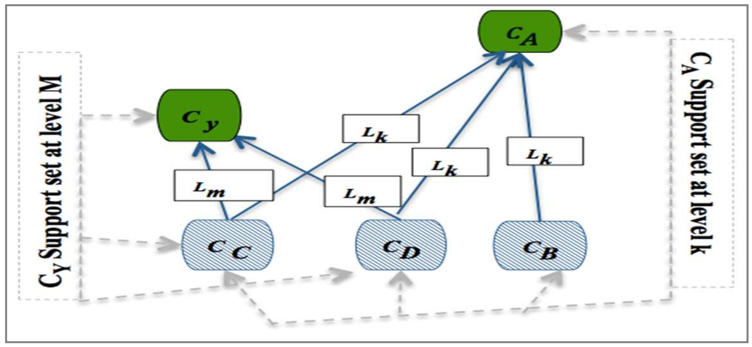


Figure 4.5 Support Node (SN) & Support Set (SS)

4.4 Zones of Derived Known Skill, DKS (k>2) & Derived Not Known Skill, DNS

Derived Known Skill (DKS) (k>2) means that there is direct evidence a learner knows a concepts C_y at a Cognitive Skill Level 2, and there is indirect evidence that he knows it at a Cognitive Skill Level higher than a Cognitive Skill Level 2. In other words, we can tell by inference that a learner could either Apply/Evaluate/Create a concept C_y. The relation condition is illustrated as the following:

If C_y is known i.e. it is in DKS (2) or VKS (2), and if all level k support nodes of C_y, i.e., S (C_y, k) is in VKS (2) $\vee$ DKS (2), then C_y will be considered as a Derived Known Skill at level k. In other words, If $C_y \in$ DKS (2) $\vee$ VKS (2) and S (C_y, k) is subset of DKS (2) $\vee$ VKS (2) $\rightarrow C_y \in$ DKS (k). Figure 4.6 illustrates DKS (k>2).

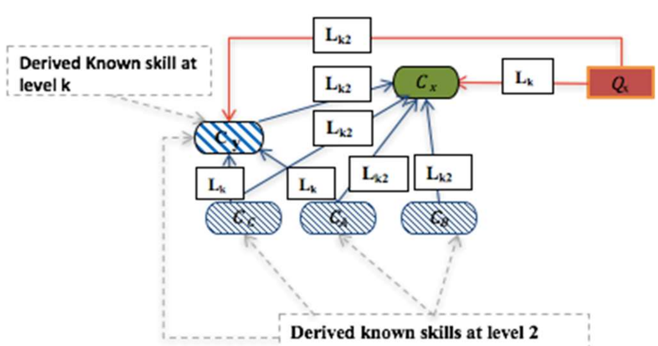


Figure 4.6 Derived Known Skills Relation at Level k

Derived Not Known Skill (DNS) (k>2) means that there is direct evidence that a learner does not know a concept C_y at Cognitive skill level 2. If this is so, we can tell by indirect evidence that he does not know the concept at a Cognitive skill level higher than a Cognitive skill level 2. The relation condition is illustrated as the following:

If C_y is known not knowing i.e. it is in DNS (2) or VNS (2), or all level k support nodes of C_y i.e. S (C_y, k) is in VNS (2) $\vee$ DNS (2) then C_y will be considered as a Derived Not Known Skill at level k. In other words, If $C_y \in$ DNS (2) $\vee$ VNS (2) or S (C_y, k) is a subset of DNS (2) $\vee$ VNS (2) $\rightarrow C_y \in$ DNS (k). Figure 4.7 illustrates DKS (k>2).

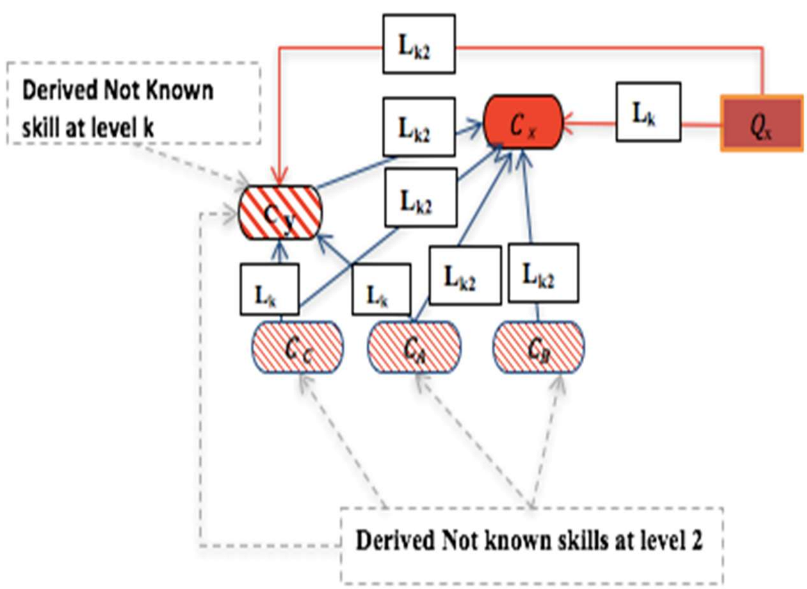


Figure 4.7 Derived Not Known Skills Relation at Level k >2

4.5 Zones of the Potential Known Skill PKS (k≥2)

Potential Known Skill (k≥2) is defined as the existence of indirect evidence that a learner knows the set of concepts that support concept A at a Cognitive Skill Level equal to or higher than 2 (Understand/Apply/Analyze/Evaluate/Create), the concept A is considered to be in the set of PKS (k≥2). The relation condition is illustrated as the following:

Let S(A, k) be the Support Set of A at level k. If every node in the S(A, k) is a subset of VKS ∨ DKS at any level (the level does not matter because we only want to guarantee that the set is known) i.e., S(A, k) ⊂ VKS ∨ DKS, but there is no evidence that A is known, then A is an element in Potential Known Skill Set PKS(k) i.e. A ∈ PKS(k) where C_d , C_x ⊂ VKS and C_C, C_A , C_B ⊂ DKS and Lk ∈ Bloom's Link at level k. Figure 4.8 illustrates the Potential Known Skills relation.

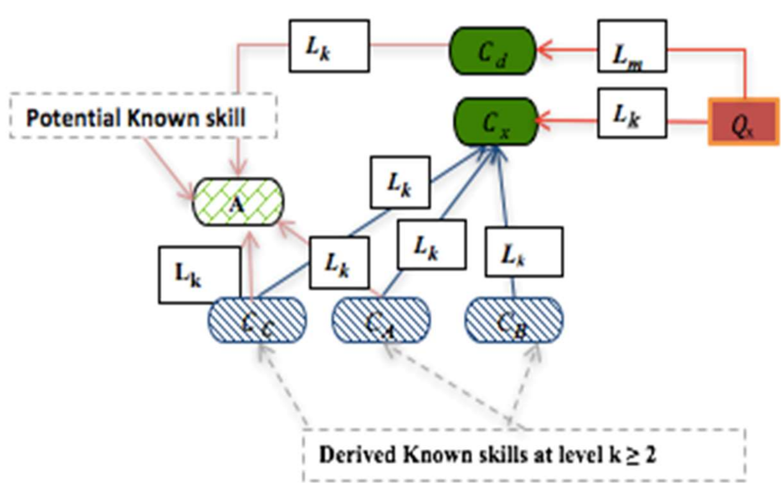


Figure 4.8 Potential Known Skill Relation

The Potential Not Known Skill at level equal to or higher than 2, PKS (k≥2), is defined as the existence of indirect evidence that a learner doesn't know concept set that support a concept A at a cognitive skills level equal or higher than 2 (Understand/Apply/Analyze/Evaluate/Create). The relation condition is illustrated as follows:

Let S(A, k) be the Support Set of A at level k. If every node in the S(A, k) is a subset of VNS ∨ DNS at any level (the level does not matter, because we only want to guarantee that the set is not known) i.e., S(A, k) ⊂ VNS∨ DNS, but there is no evidence that A is known or not known, then A is an element in the Potential Not Known Skill set PNS(k) i.e. A ∈ PNS(k), where C_d , C_x ⊂ VNS and C_C, C_A , C_B ⊂ DNS and Lk ∈ Bloom's link at level k. Figure 4.9 illustrates Potential Not Knowing Skill relation.

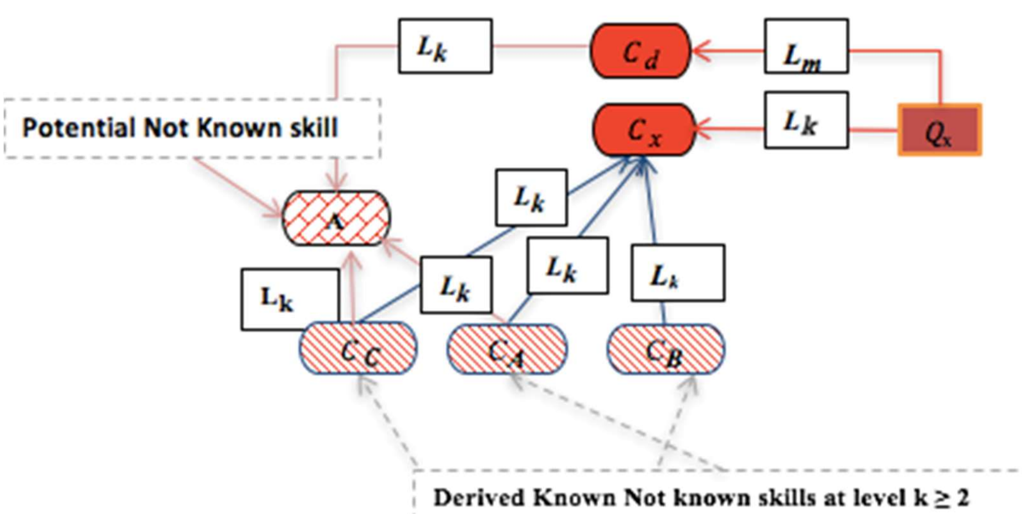


Figure 4.9 Potential Not Known Skill Relation

4.6 The Assessment Analytics of the TCS^2 Theory

In this study, a new parameter, Cognitive Skill Level, was added to the knowledge assessment. Therefore, we provided Assessment Analytics to the TCS^2 theory. The Assessment Analytics is an assessment analysis method to connect the concepts in the test for CLMCG. The Cognitive Skill Levels refer to levels such as whether a learner has acquired the concept at the level of Understanding, Applying, Analyzing, Evaluating, or Creating. Providing a knowledge assessment analysis method helps design a proper test which can measure exactly the covered knowledge of the course objective. The results of the Assessment Analytics are the Concept States that classify the set of the tested concepts into two sets, known and not known, which inform whether a learner has already learned or hasn't learned, is ready to learn, or is not ready to learn a concept at certain skill level. The assessment analysis maps the proposed CLMCG, and the sets of Concept States to estimate the Concept States from the learner answers.

4.6.1 Connecting the Questions to the CLMCG.

A test usually comprises a set of questions. Answering question Q requires knowledge about a set of concepts. A simple question is designed to measure the pedagogical skill of the learner targeting a concept. Thus, question Q may be linked to a set of concepts, in which each link could be labeled according to the appropriate skill level. The proposed test could be connected to the concept map CLMCG.

When a learner successfully answers a question, it may be concluded that he/she has learned each concept associated with the question at a certain skill level.

As answering a question Q requires pedagogical knowledge about a set of concepts, the question Q can be linked to a set of concepts, of which, each link could be labeled according to the appropriate skill verb of RBT. A relationship example to illustrate this would be: to answer a question Q_i correctly, a learner should know the concept C_x at skill level Lk.. Thus, the link between question Q_i and concept C_x is labeled with the skill verb Lk. Accordingly, when the learner successfully answers a question, one could conclude that he/she has learned the concepts connected to the question at a certain skill level.

4.6.2 The Process to Estimate the Concepts States

The first step is to connect the question with CLMCG. A test normally is composed of a set of questions. Answering question Q requires knowledge about a set of concepts. The grader determines the exact skill that is associated with each tested concept. Thus, question Q could be linked to a set of concepts in CLMCG, and each link could be labeled according to the tested skills.

The second step is the estimation of VKS. When a learner successfully answers a question, one could conclude that he has learned the concepts associated with the question at a certain tested skill. Thus, each concept in the question may be assigned to a certain skill level. The concepts, which are directly tested by the question and correctly answered by the learner, form his Concept State of VKS.

Third, to identify the untested concept set DKS, an expert should determine the Support Concept Set SS of VKS at the exact tested level. Thus, to be able to answer the VKS, the learner must know the concept set at a certain level. This concept set is the DKS. If the learner answers the VKS correctly, it may be estimated that he/she knows the related DS. Only the concept set supporting the correct answer is added to the learner Concept State of DKS.

Fourth, for extracting PKS, the expert needs to determine the concept set for which most of the SSs at a certain level are in VKS or DKS. Only the correct answer of VKS or DKS is considered to indicate the Concept State of PKS to be added to the learner knowledge.

Fifth, repeat the three steps for each concept in the test.

Sixth, each question in the test will be connected to CLMCG as the appropriate VKS, DKS and PKS.

Seventh, to represent the Concept State of the learners in the graph, each learner's answer will be mapped to the CLMCG, and each learner will get her/his own estimation set of the Assessment TCS^2 theory. If the detected concept set already existed in the CLMCG, the evaluator just needs to connect the question to the related concepts in the graph. Next, the learner's answer to the question will connect to the mapped concepts in the graph according to the presented states. Thus, we can estimate the learners' skills based upon the three states: VKS, DKS and PKS. Accordingly, when a learner successfully answers a question, one can conclude what skills he has learned and what he is ready to learn from the mapped graph.

4.6.3 Example of an Assessment Scenario

- A preliminary test of the course is given to the learner.
- The preliminary test contains the basic learned concepts, which are also included in the course objective domain. They may have already been tested during the course in quizzes. Thus, the preliminary exam is an accumulation of the total of all of these quizzes.
- The assessment analysis will then be processed on the learner results.
- The grader must correct the answer based on the Cognitive Skill Level of each tested concept, according to the Assessment Analytics of the Theory CS^2.
- The result of the assessment analysis provides the true picture of the concepts attained by the learner at a certain level to achieve the learning objective.
- The result of the learner assessment reveals the knowing and unknowing concepts at a certain level in terms of the Concept States.
- The Concept States are VKS, VNS, DKS, DNS, PKS and PNS.
- Once the unknowing Concept States are known, this leads to another post-test to reveal the prerequisite concepts which caused the failure to attain the concepts at the tested skill level of the learning objectives.
- If a post-test is conducted and the true unknown concepts forming the DNS are realized, then the final learner profile will contain three sets of concepts

combined with the skill levels: known concepts at skill levels, unknown concepts at skill levels, ready to be known and not ready to be known concepts at skill levels by probability.

- If there is a contradiction between the concepts in the knowing and not knowing sets, then the probability computation will be applied to realize precise Knowledge States estimation.
- The output of the assessment is introduced to the learner.

4.6.4 Example of Question to Illustrate the Concept States

This section provides an example of a question that illustrates the Theory of Cognitive Skill in Concept Space (TCS^2). One question is used and the analytical steps are applied to it to extract the three sets

Example 4.1. Suppose a test question is given for learners in algorithm course such as:

Q_1 Show the order of elements in the [given] array after each pass of the Selection Sort Algorithm. int arr[6] = {5, 1, 8, 2, 7, 9}   [write the final result in the array]

The Assessment Analytics are interpreted with the question to conclude the cognitive relation.

The tested concepts set are TC = {“Selection Sort Algorithm”, “Sort”, “Array”, “The Order”, “The Passes”}. The objective tested concept is “Selection Sort Algorithm.” In conventional evaluation, a grader will assign a quantitative score for the learner based on the answer of the objective tested concept.

Let us highlight the cognitive relation on the tested concepts. To answer question Q_1 correctly, the concepts "Sort" and "The Order" need to be understood, the concept "Array" needs to be applied, the concept "The Order" of the element in "the Array" needs to be applied, "The Passes" of "Selection Sort Algorithm" need to be evaluated and applied and the concept "Selection Sort Algorithm" needs to be applied. The TCS^2 Assessment Analytics will be applied to the tested concepts to estimate the skill levels of the Concept States as the introduced conditions of VKS, DKS and PKS.

The Verified Skills are as follows:

1. The Verified Skills set at level 2 is VKS (2) = {"Order", "Set"}
2. The Verified Skills set at level 3 is VKS (3) = {"Selection Sort Algorithm", "The Passes of Selection Sort Algorithm", "Array", "The Order of the Element in the Array"}
3. The Verified Skills set at level 4 VKS (4) = {"The Order of the Element in the Array"}
4. The Verified Skills set at level 5 is VKS (5) = {"The Passes of Selection Sort Algorithm"}
5. The Derived skills set will be as following:
6. DS (2) = {"Selection Sort", "Sort Process", "The Algorithm", "Unsorted Order", "The Places Step of Selection Search Algorithm", "Illustration example of the Selection Sort Algorithm", "Selection Sort Function", "The Result List", "Ascending Order", "The Smallest Element", "The Largest Element", "The Elements" , "The Content" , "Iteration", "The Places Step of Selection Sort", "

Traversal of the Elements”, “Position”, “Process”, “Places”, “Step”, “Selection”, “Search”, “Illustration”, “Example”, “Ascending”, “Function”, “Result”, “Smallest”, “Largest”, “Sort”}

7. DS (3) = {“Selection Sort”, “The Result List”, “Ascending Order”, “The Largest Element”, “The Smallest Element”, “Traversal of the Elements”, “Iterations”, “The Places Step”, “Sort Process”, “Passes”, The Resulting List”}

The Potential Skill Set will be:

1. PKS (3) = {“The Index of Smallest Element”, “Sublist Array”, “Simple Search Algorithm”, “Radix Sort Algorithm”, “Heap Sort Algorithm", “Insertion Sort Algorithm”, “Sort”}
2. PKS (2) = {“Analyses of the Algorithm”, “Algorithm Performance”, “Running Time”, “Selection Sort Running Time $O(n^2)$”}

If the learner answers the entire concepts in the question correctly then one can estimate his VKS, DKS, and PKS as pointed in the Assessment Analytics of the TCS^2 theory.

CHAPTER 5

The Probability Computation to Estimate the Concept States of the Learners

5.1 Introduction

In this chapter, we discuss the calculation of the probability of knowing the concepts in the Concept Space. In chapter 4, we provided the theories to identify the internal relationship between every concept, in either single word form or phrase form. In this chapter, we identify the numerical values of each concept in the concept states, which have been introduced in chapter 4. Also, in this chapter, we study the difficulty in calculating the probability of knowing or not knowing the concepts in conditions of contradiction and in conditions of the concepts relations. Subsequently, the chapter starts with a problem statement; and goes onto suggest solutions and provide illustrated examples. In the example we study, we use Bayes' Theorem to find the probability of knowing the concepts, and we assume that there is no dependency cycle existing in the graphs of the relation between the concepts. For example, there is a dependency between the concepts A, B, C and D but there is no repeating cycle. (D doesn't go back to A).

5.2 The Problem Statement

Given a set of questions Q = $\{q_1, q_2, q_3, \dots q_n\}$ and a set of evaluations E = $\{e_{q_1}, e_{q_2}, e_{q_3}, \dots . e_{q_n}\}$.

Let P(C|R) be the conditional probability of knowing a concept C using the given set of responses R.

Let m_r , g_r = 0.2. (Arbitrary assumption for test the proposed questions). Therefore, the probability of knowing the concept C is $\mathrm{P(C|Q_r)}$= 0.8 and the probability of not knowing the concept $\bar{\mathrm{C}}$ is $\mathrm{P(\bar{C}|Q_r)}$= 0.2.

Let a learner answer a question q, which gives evidence of his/her state of knowledge about a concept C. If the response is correct, then the probability of knowing the concept C is $\mathrm{P(C|}Q_\mathrm{r})$ $= \mathrm{P(Q_r|C)} = (1 - g_r)$ and the probability of not knowing the concept C is $\mathrm{P(\bar{C}|Q_r)} = \mathrm{P(Q_r|\bar{C})}$= g_r. On the other hand, if the response to a question q is incorrect then the probability of knowing the concept C, which has been asked by the question $\mathrm{q_r}$ is $\mathrm{P(C|\bar{Q}_r)}$ = $\mathrm{P(\bar{Q}_r|}m_r$) = m_r, and the probability of not knowing the concept C is $\mathrm{P(\bar{C}|\bar{Q}_r)}$ = $\mathrm{P(\bar{Q}_r|\bar{C})}$ = (1−m_r). The two constants, m_r , g_r ∈ $[0, 1)$, are respectively called (careless) error probability and guessing probability at q_r. The subscript r indicates index of the question number. What is the probability of knowing the concept C if the set of evaluations E $= \{\mathrm{e_{q_1}, e_{q_2}, e_{q_3}, \dots\ e_{q_n}}\}$ is known for the responses to the set of questions Q = $\{q_1, q_2, q_3, \dots\ q_n\}$?

5.3 The Proposed Solution Based on Bayes' Theorem:

Intuitively, it is suggested that Bayes' Theorem could be used to calculate the probability of knowing a concept, even though the concept is evaluated based on reflected evaluations of the concept. Also, it could be used to calculate the probability of knowing the concept, even in the existence of complex relation between the concepts in the Concept Space, such as the relation between the concepts in the of VS, DS and PS.

5.3.1 Introduction to Bayes' Theorem

There are many versions of Bayes' Theorem[10]. In this study, we investigated two versions. The first version is the basic formula, which is a simple form used to calculate a probability of an event based on conditions that might be related to the event. The second version is the extended form of the simple formula of Bayes' Theorem, which is used to calculate probability of an event based on many conditional events observed and affect the probability of the evaluated event.

1. **The basic Bayes' Theorem**

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \quad ,$$

where A and B are events and $P(B) \neq 0$.

- P(A) and P(B) are the probabilities of observing A and B without regard to each other. It is the initial degree of belief in A or B.
- P(A|B), a conditional probability, is the probability of observing event A given that B is true.
- P(B|A) is the probability of observing event B given that A is true.

In this equation, we could calculate the probability of A given that B is true. There is only one observation of a single event B. Here, B is a single event observed and it affects the probability of the related event A. What if the probability of A is based on many observations? Or in various possible events A is such that A & B could occur

[10] http://homepages.wmich.edu/~mcgrew/Bayes8.pdf

together in various possibilities?[11] Or, if there are many conditional events, which affect the probability of the related event A? This leads us to use the extended form of Bayes' Theorem.

2. The Extended Formula of Bayes' Theorem

In the case where there are many observations (evidences or references) that indicate the knowing of the concept, then the extended form of Bayes' Theorem is used. Moreover, the observations, which indicate the knowing of the concept, are reflected. For these reasons, we should use Bayes' Theorem in the extended form to find out the correct value of the probability.

The extended formula of Bayes' Theorem is generally encountered when looking at two competing statements or hypotheses[12] such as, correct and incorrect, or knowing and not knowing. In other words, we consider the impact of B having been observed on our belief in various possible events A. In our work, the possibilities are that the concept C is either known or unknown.

Extended formula of Bayes' Theorem is

$$\mathrm{P(A|B)} = \frac{\mathrm{P(B|A)} * \mathrm{P(A)}}{\mathrm{P(B|A)} * \mathrm{P(A)} + \mathrm{P(B|\overline{A})} * \mathrm{P(\overline{A})}}$$

[11] In many applications, for instance in Bayesian inference, the event B is fixed in the discussion, and we wish to consider the impact of its having been observed, on our belief, in various possible events A. In such a situation, the denominator of the last expression, the probability of the given evidence B, is fixed; what we want to vary is A.

[12] https://en.wikipedia.org/wiki/Bayes (McGrew, 2005)%27_theore

in the point extended form

- $P(A)$, the probability A, is the prior probability or the current unconditional probability, which is the initial degree of belief in A.
- $P(\overline{A})$, is the corresponding probability of the initial degree of belief against A: $1 - P(A) = P(\overline{A})$
- $P(B|A)$, the conditional probability or likelihood, is the degree of belief in B, given that the proposition A is true.
- $P(B|\overline{A})$, the conditional probability or likelihood, is the degree of belief in B, given that the proposition A is false.
- $P(A|B)$, the posterior probability, is the probability for A after considering B for and against A.

Based on the given problem statement, we suggest the extended formula of Bayes' Theorem to be used. Example 5.1 illustrates the solution.

3. The Law of Total Probability

The Law of Total Probability is used in Bayes' Theorem. The Law of Total Probability is identified as the Theorem (3.17) in point 3.3.4 of the section of probability theorems in the research of Zwillinger & Kokoska (2000, p. 41). Inferred from the work of McGrew (2005), I elucidated that the extended version of Bayes' Theorem is derived from the Law of Total Probability. The Law of Total Probability was significant in extending Bayes' Theorem, wherein the Law of Total Probability could be replaced in the denominator of Bayes' Theorem.

The Theorem of the Law of Total Probability is such that:

Suppose $A_1, A_2, A_3, \ldots, A_n$ is a collection of mutually exclusive, exhaustive events,

$P(A) \neq 0$, i = 1,2,…..,n.

For any event B:

$P(B) = \sum_{i=1}^{n} P(B|A_i).P(A_i)$ Theorem 1

The Theorem 1 according to the given information in the problem statement: Let B be the event of the responses data R,

While A is the set of collection events of the concept C. There are only two mutual exclusive events A_1 and A_2, where A_1 is the set of the events of knowing the concept C, A_2 is the set of events of not knowing the concept.

Thus,

$P(R) = \sum_{i=1}^{n} P(R|A_i).P(A_i)$

$P(R) = P(R|A_1).P(A_1) + P(R|A_2).P(A_2)$ Theorem 1.1

It is known A_2 is a complement set of A_1

Thus, the extended formula of Bayes' Theorem is concluded by replacing Theorem 1.1 in the denominator of the basic Bayes' Theorem.

*As observed the probability Theorem 1.1 is a special case of the Law of Total Probability, Theorem 1.

5.4 Examples to Prove the Suggested Solution by Using Bayes' Theorem

Bayes' Theorem could be used to calculate the probability of knowing the concept C in two cases:

1. If we have only one response to a question asked about the concept, then we can use the basic form, which is the basic Bayes' Theorem.

2. If we have information about the concept C such as a related set of concepts or set of questions, then we can use the extended Bayes' Theorem by calculating the denominator value to include all the cases that indicate knowing the concept C.

The reason I use the extended form is that it makes possible to compute the probability of an event B, where B consists of several observations. For following reasons: we assume that the learner either knows or doesn't know the concept C, the probability of a given set of responses depends on whether the learner knows or does not know the concept. The extended form of Bayes' Theorem is used.

5.4.1 Example 5.1: The Probability of Knowing a Concept Evaluated by More than One Question

This example introduces the cases of one concept evaluated by many questions. The challenge here is that the questions could also include conflicted evaluations of the concept.

Suppose that a set of questions Q_q asks about the concept C_j at skill level k, which denoted as C_j^{Lk}. The set of questions is $Q_q = \{q_1, q_2, q_3, q_4, q_5, q_6\}$. The set of responses $R = \{Q_1, Q_2, \overline{Q}_3, Q_4, Q_5, \overline{Q}_6\}$. Let m_r , g_r = 0.2. (Arbitrary assumption for testing the proposed questions). If the response is correct, then the probability of knowing the concept C is $P(C|Q_r) = P(Q_r|C) = (1 - g_r)$ and the probability of not knowing the concept C is $P(\overline{C}|Q_r) = P(Q_r|\overline{C}) = g_r$. On the other hand, if the response to a question q is incorrect then the probability of knowing the concept C, which has been asked by the question q_r is $P(C|\overline{Q}_r) = P(\overline{Q}_r|m_r) = m_r$ and the probability of not knowing the concept C is $P(\bar{C}|\bar{Q}_r) = P(\bar{Q}_r|\bar{C}) = (1-m_r)$. Therefore, the probability of knowing the concept C is $P(C|Q_r)$= 0.8 and the probability of not knowing the concept $\bar{C}$ is $P(\bar{C}|Q_r)$= 0.2. Figure 5.1 illustrates the Example 5.1.

The question is: given the evidence R, find the probability of knowing the concept $P(C_j^{Lk}|R)$.

The next section is to explain the computation of $P(C_j^{Lk}|R)$ by using Baye's Theorem. The result of the example is illustrated in Tables 5.1.

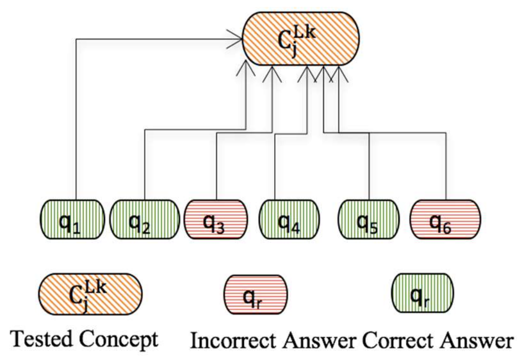


Figure 5.1 Many Questions Asked About One Concept Illustration of Example 5.1

5.4.2 The Basic Formula of Bayes' Theorem

By using the first formula of Bayes, which is the basic formula, we should calculate the probability of knowing the concept C_j^{Lk}, on condition of the information of all the related concepts.

So, how can we use the first formula?

The basic Bayes' Formula allows for calculating the probability of an event in the condition of observing previous event. If we calculate the condition of the related concepts one by one, then could the result be acceptable?

Let's try it:

From basic Bayes' Theorem, the Equation 1 is:

$$\mathbf{P}\left(\mathbf{C}_{\mathbf{j}}^{\mathbf{Lk}} \middle| \mathbf{Q}_{\mathbf{r}}\right) = \frac{\mathbf{P}\left(\mathbf{Q}_{\mathbf{r}} \middle| \mathbf{C}_{\mathbf{j}}^{\mathbf{Lk}}\right) * \mathbf{P}\left(\mathbf{C}_{\mathbf{j}}^{\mathbf{Lk}}\right)}{\mathbf{P}(\mathbf{Q}_{\mathbf{r}})}$$

- C_j^{Lk} is the concept C_j at skill level Lk.

- Q_r is a correct response to the question q_r

- $P(C_j^{Lk})$ is the unconditional probability, which is the probability of observing knowing the concept C_j^{Lk} without regard for observation of another concept or question.

- $P(C_j^{Lk}|Q_r)$ is a conditional probability, the probability of knowing the concept C_j^{Lk} given the observing response Q_r to the question q_r.

- $P(Q_r|C_j^{Lk})$ is a conditional probability, the probability of observing correct response to the question q_r on the condition the event of knowing concept C_j^{Lk} is true.

The fake solution:

Let's start by finding the probability of knowing the concept C_j at skill level Lk denoted as C_j^{Lk}.

The unconditional probability of knowing a concept C_j^{Lk} is $P(C_j^{Lk}) = 0.5$ from Equation 5.

$P(Q_1|C_j^{Lk}) = 0.8$ from the definition of the probability of correct response to a question q_1. This is an assumption that the probability of knowing the concept is known from response to question q_1.

Let us choose two questions asked about the concept C_j^{Lk} with conflicted responses. From Figure 5.1, question q_1 gets correct responses by a learner, whereas question q_6 gets wrong response by the same learner. The question now is that on the condition of which of them can we calculate the conditional probability of knowing concept C_j^{LK}?

Let $P(Q_1) = 1$ and $P(\overline{Q}_1) = 0$.

$$P(C_j^{Lk}|Q_1) = \frac{P(Q_1|C_j^{Lk}) * P(C_j^{Lk})}{P(Q_1)} \qquad = \frac{0.8 * 0.5}{1} = 0.4 \text{ is it true??}$$

We cannot tell if it is true, since there is another response asked about the concept C_j^{Lk}, which affects the probability of knowing the concept C_j^{Lk}. Moreover, the probability of the affecting concept is present in both the numerator and the denominator and cancels upon simplification. We observed that, $P(C_j^{Lk}) = P(C_j^{Lk}|Q_1)$. We can tell that the event of knowing the concept C_j^{Lk} wasn't affected by the event of a response to the question asked about C_j^{Lk}. Therefore, using basic formula of Bayes' Theorem is an appropriate solution to calculate the probability of knowing a concept based on one observation. Moreover, if we have only one response as an evidence, then $P(C_j^{Lk})$ should be equal to 1, since the answer is correct and $P(C_j^{Lk}|Q_1) = P(Q_1|C_j^{Lk}) = 0.8$.

Let us calculate the probability of knowing the concept C_j^{Lk} on the condition of incorrect answer such as the response to the question q_6:

$$P(C_j^{Lk}|\bar{Q}_6) = \frac{P\left(\bar{Q}_6|C_j^{Lk}\right)*P(C_j^{Lk})}{P(\bar{Q}_6)} = \frac{0.2*0.5}{0} = \text{invalid}$$

It is obvious that we cannot use Bayes' basic formula to calculate the probability of knowing the concept C_j^{Lk} since we should use only one observed instance. Therefore, if we have many observed instances, or a set of data, indicating the "knowing" of the concept C_j^{Lk}, then we have to use the extended formula of Bayes' Theorem.

5.4.3 Using the Extended Formula of Bayes' Theorem

If there is previous information based on a set of related concepts which provide evidences about the evaluated concept C_j^{Lk}, then the formula of Bayes' Theorem to be used is the Extended formula of Bayes' Theorem.

In the same way, if we have a set of questions q giving some evidence about a concept C_j^{LK} with a contradiction, like some of them give evidence that the concept is known and others that it is not known, then the appropriate Bayes' Formula to use will be the extended formula of Bayes (Equation 2) [13] (Lee, 2012).

From Bayes' Theorem, the Equation 2 is

$$P(\mathbf{C_j^{Lk}}|R) = \frac{P\left(\mathbf{R}\middle|\mathbf{C_j^{Lk}}\right)*P(C_j^{Lk})}{P\left(\mathbf{R}\middle|\mathbf{C_j^{Lk}}\right)*P\left(C_j^{Lk}\right)+P\left(\mathbf{R}\middle|\mathbf{\bar{C}_j^{Lk}}\right)*P\left(\bar{C}_j^{Lk}\right)} \quad ,$$

[13] https://en.wikipedia.org/wiki/Bayes%27_theorem

where

- C_j^{Lk} denotes knowing the concept C_j^{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}. I suggest the ideal value of $P(C_j^{Lk})$ to be the ratio number of the correct responses to the number of questions asked about the concept C_j^{Lk}.

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

Thus, Equation 3 is

$$P(\mathbf{C_j^{Lk}}) = \frac{|\mathbf{Q}|}{|\mathbf{Q}_q|} \ ,$$

where Q is the set of the correct responses to the questions asked about the concept C_j^{Lk}

Q_q is the total number of questions asked about the concept.

$P(\bar{C}_j^{Lk})$ 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}. The equation to calculate it is Equation 4

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

Thus, Equation 4 is

$$P(\bar{\mathbf{C}}_\mathbf{j}^{\mathbf{Lk}}) = \frac{|\bar{Q}|}{|Q_q|} \ ,$$

where $\bar{Q}$ is the set of the incorrect responses to the questions asked about the concept C_j^{Lk}.

If there is no previous information about the concept C_j^{Lk}, this means that no question has been asked about the concept before. Therefore, I assume that the probability of knowing and not knowing to be equally likely, i.e.,

$P(C_j^{Lk}) = \frac{1}{2}$ Equation 5

If we have information about the concept such that, the probability of knowing a concept C_j^{Lk}, by given a set of data, $P(C_j^{Lk}|R)$ is already calculated and a new question in new session is asked about the concept C_j^{Lk}, then the initial unconditional probability $P(C_j^{Lk})$ will be replaced with $P(C_j^{Lk}| E_{i-1})$, were E_{i-1} is previous probability information about C_j^{Lk}.

Thus, $P(C_j^{Lk}| E_i)$ is replaced by $P(C_j^{Lk}| E_{i-1})$.

- $P(R|C_j^{Lk})$ 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}, where we defined it based on Equation 6, which is illustrated later in section 5.4.4.

- $P(R|C_j^{Lk})$ is the conditional probability of the event that the responses in R occur, conditional on the event of not knowing the concept C_j^{Lk}, where we defined it based on Equation 6 which is illustrated later in section 5.4.4.

5.4.4 The Computation of $P\left(R\middle|C_j^{Lk}\right)$ & $P\left(R\middle|\bar{C}_j^{Lk}\right)$ (The Products)

$P(R|C_j^{Lk})$ is the conditional probability of the event that the responses in R occurs, that means conditional independence on the event of knowing the concept C_j^{Lk}. In other words, the response to the question q_r denoted by Q_r is conditional independence on the concept C_j^{Lk}. It is computed using Equation 6.

The Equation 6 is

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

where $P(R|C_j^{LK}) = \prod_{r=1}^{n} P(Q_r|C_j^{LK}) =$ The multiplication of the response data given knowing the concept C_j^{LK}. C_j^{LK} is the concept C_j at skill level Lk. Q_r is the response to the question q_r;

Thus,

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

This multiplication of the conditional independent events is called multiplication rule for independent events. If the response to a question Q_r is correct, then the probability of knowing the concept C_j^{LK}, which has been asked by the question q_r, is $P(Q_r|C_j^{LK}) =$ *$(1-g_r)$*. On the other hand, if the response to a question q is incorrect, then the probability of knowing the concept C_j^{LK}, which has been asked by the question q_r, is $P(\overline{Q}_r|C_j^{LK}) = m_r$. The two constants, $m_r, g_r \in [0,1[$, respectively called (careless) error probability and guessing probability at q_r, "r" is an index refers to the question number, 1,2,….. of the related question to the evaluated concept.

$P(Q_r|\overline{C}_j^{LK}) =$ is the conditional probability of the event that the responses in R occurs, conditional on the event of not knowing the concept C_j^{LK}. It is calculated using Equation 7.

The Equation 7 is:

$\mathrm{P}(\mathrm{R}|\bar{\mathrm{C}}_j^{LK}) = \prod_{r=1}^{n} \mathrm{P}\,(\mathrm{Q}_r|\bar{\mathrm{C}}_j^{LK})$,

where

$\mathrm{P}(\mathrm{R}|\bar{\mathrm{C}}_j^{LK}) = \prod_{r=1}^{n} \mathrm{P}\left(\mathrm{Q}_r\middle|\bar{\mathrm{C}}_j^{LK}\right) =$ The multiplication of the conditional independent response data given not knowing the concept $\bar{\mathrm{C}}_j^{LK}$. Thus,

$\prod_{r=1}^{n} \mathrm{P}\left(\mathrm{Q}_r\middle|\bar{\mathrm{C}}_j^{LK}\right) = \{\mathrm{P}(\mathrm{Q}_1|\bar{\mathrm{C}}_j^{LK}) * \mathrm{P}(\mathrm{Q}_2|\bar{\mathrm{C}}_j^{LK}), * \ldots \mathrm{P}(\mathrm{Q}_n|\bar{\mathrm{C}}_j^{LK})\}$,

where $\mathrm{P}\left(\bar{\mathrm{Q}}_r\middle|\bar{\mathrm{C}}_j^{LK}\right) = (1-m_r)$ for wrong response to q_r and $\mathrm{P}\left(\mathrm{Q}_r\middle|\bar{\mathrm{C}}_j^{LK}\right) = g_r$ for correct response to q_r.

5.4.5 The Conditional Probability of Not Knowing the Concept C_j^{Lk} Given the Data of Set of Responses

We proved the result of the proposed equations by using Bayes' Theorem to calculate the conditional probability of not knowing the concept $\bar{\mathrm{C}}_j^{LK}$ by given the set of responses $\mathrm{P}(\bar{\mathrm{C}}_j^{LK}|\mathrm{R})$ and compared the result with the result of $1-\mathrm{P}(\mathrm{C}_j^{Lk}|\mathrm{R})$.

From the extended formula of Bayes' Theorem, the equation to calculate $\mathrm{P}(\bar{\mathrm{C}}_j^{LK}|\mathrm{R})$ is Equation 8.

The Equation 8 is

$$\mathrm{P}(\bar{\mathrm{C}}_j^{LK}|\mathrm{R}) = \frac{\mathrm{P}\left(\mathrm{R}\middle|\bar{\mathrm{C}}_j^{LK}\right) * \mathrm{P}\left(\bar{\mathrm{C}}_j^{LK}\right)}{\mathrm{P}\left(\mathrm{R}\middle|\bar{\mathrm{C}}_j^{LK}\right) * \mathrm{P}\left(\bar{\mathrm{C}}_j^{LK}\right) + \mathrm{P}\left(\mathrm{R}\middle|\mathrm{C}_j^{Lk}\right) * \mathrm{P}\left(\mathrm{C}_j^{Lk}\right)}$$

5.4.6 The Solution by Using the Extended Formula of Bayes' Theorem

Example 5.1 illustrates the evaluation of the tested concept by many questions. The related questions which asked about the concept could have reflected responses, even though they asked about the same skill level of the concept. In this case, the solution of Example 5.1 would obtained by using the extended formula of Bayes' Theorem

The question was, suppose a set of questions $Q_q = \{q_1, q_2, q_3, q_4, q_5, q_6\}$ are directly or by inference asked about the concepts. The set of responses to the questions is $R = \{Q_1, Q_2, \overline{Q}_3, Q_4, Q_5, \overline{Q}_6\}$. The total number of questions = $|Q_q| = 6$.

The responses of the learner as the following:

-The set of correct responses $R_Q = \{Q_1, Q_2, Q_4, Q_5\}$ therefore $|R_Q| = 4$.

-The set of wrong responses $R_{\overline{Q}} = \{\overline{Q}_3, \overline{Q}_6\}$ therefore $|R_{\overline{Q}}| = 2$.

Thus, $P(C_j^{LK}) = \frac{4}{6} = \frac{2}{3} = 0.67$ From the Equation 3

5.4.7 The Multiplication of the Probabilities

$\prod_{r=1}^{n} P(R|C_j^{LK}) = \{P(Q_1|C_j^{LK}) * P(Q_2|C_j^{LK}) * P(\overline{Q}_3|C_j^{LK}) * P(Q_4|C_j^{LK}) * P(Q_5|C_j^{LK}) * P(\overline{Q}_6|C_j^{LK})\}$. From Equation 6

I suppose probabilities of knowing a concept with assumption of errors such that $P(Q_r|C_j^{LK}) = (1-g_r)$ for a correct answer and $P(\overline{Q}_3|C_j^{LK}) = m_r$ for a wrong answer. "m_r", "g_r" $\in [0,1[$, are respectively called (careless) error probability and guessing probability at q_r. "r" refers to the number of the related question. In multiple choice questions the error probability is high.

Let m_r, g_r = 0.2. (arbitrary assumption for testing the proposed equations). Therefore, $P(Q_r|C_j^{LK}) = 0.8$ for knowing the concept and $P(Q_r|\overline{C}_j^{LK}) = 0.2$ for not knowing the concept.

The conditional probability of the set of the responses given knowing the concept C_j^{LK} is

$P\left(R \middle| C_j^{LK}\right) = 0.8 * 0.8 * 0.2 * 0.8 * 0.8 * 0.2 = \ 0.016$ From Equation 6

$P\left(R \middle| \bar{C}_j^{LK}\right) = 0.2 * 0.2 * 0.8 * 0.2 * 0.2 * 0.8 = \ 0.001$

The calculation of knowing a concept C_j^{LK} given a set of evaluations by using Bayes' Theorem, Equation 2

$$P(C_j^{LK}|R) = \frac{0.016 * 0.67}{0.016 * 0.67 + 0.001 * 0.33}$$

$$\frac{0.011}{0.01 \quad 0.00033} \qquad = \frac{0.011}{0.011}$$

Replacing the unconditional probability $P(C_j^{LK})$ and $P(\bar{C}_j^{LK})$ in Equation 2

$P(C_j^{LK}|R) = 1$

The conditional probability of not knowing the concept C_j^{LK}, $P(\bar{C}_j^{LK}|R) = 1 - 1 = 0$

Also, we proved the result by using Bayes' Theorem to calculate the conditional probability of not knowing the concept C_j^{LK} by the given set of responses R. We proved the correction of Equation 2 we can use Equation 8 to calculate the probability of not knowing the concept C_j^{LK} and compare it with the result of $1 - P(\bar{C}_j^{LK}|R)$.

Also, by using Bayes' Theorem, as Equation 8

$$P(\bar{C}_j^{LK}|R) = \frac{P\left(R \middle| \bar{C}_j^{LK}\right) * P\left(\bar{C}_j^{LK}\right)}{P\left(R \middle| \bar{C}_j^{LK}\right) * P\left(\bar{C}_j^{LK}\right) + P\left(R \middle| C_j^{LK}\right) * P\left(C_j^{LK}\right)} = 0.$$

The probability of not knowing the concept C_j^{LK}, $P(\bar{C}_j^{LK}|R) = 1 - 1 = 0$.

Thus, we proved the equations which used Bayes' Theorem to calculate the conditional probability of not knowing the concept C_j^{LK}, given the set of responses R. Particularly, we proved Equation 6 by calculating the probability of the concept is unknown. Then, we compared the two results, one by using Equation 8 to replace Equation 7, and another as result of the conditional probability of not knowing the concept by the cumulative equation:

$P(\bar{C}_j^{LK}|R) = 1 - P(C_j^{LK}|R)$.

$P(\bar{C}_j^{LK}|R) = \frac{P(R|\bar{C}_j^{LK}) * P(\bar{C}_j^{LK})}{P(R|\bar{C}_j^{LK}) * P(\bar{C}_j^{LK}) + P(R|C_j^{LK}) * P(C_j^{LK})} = 0$ By using Equation 8.

The example result is in tables 5.1 and 5.2

Table 5.2 The Probability of Knowing the Concept with Conflicted Evaluations in Example 5.1

R_Q	R	$P(C_j^{LK})$ Eq3	$P(R\|C_j^{LK})$ Eq6	$P(C_j^{LK}\|R)$ Eq2
4	{0.8,0.8,0.2,0.8,0.8 ,0.2}	4/6= 0.67	0.8*0.8*0.2*0.8*0.8*0.2 = 0.016	1

Table 5.1 The Probability of Not Knowing the Concept with Conflicted Evaluations in Example 5 1

$R_{\bar{Q}}$	R	$P(\bar{C}_j^{LK})$ Eq4	$P(R\|\bar{C}_j^{LK})$ Eq7	$P(\bar{C}_j^{LK}\|R$ Eq8
2	{0.8,0.8,0.2,0.8,0.8, 0.2}	1- 0.67= 0.33	0.2*0.2*0.8*0.2*0.2*0.8 = 0.001	0

5.4.8 Example 5.2: Illustration of the Calculation of the Probability of Knowing the Concepts at Skill Levels Based on the Proposed Methods

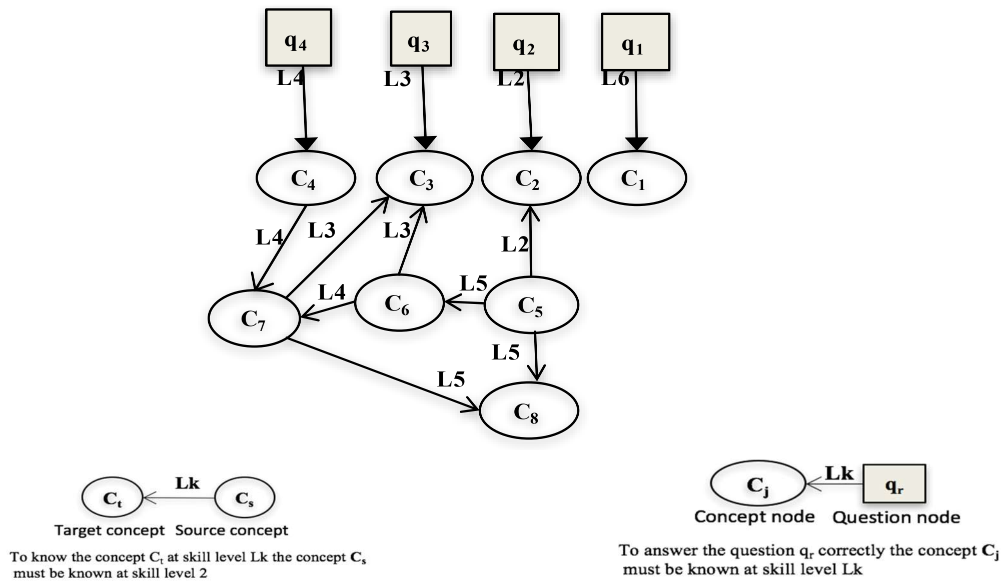


Figure 5.2 The Structure of the Concept Relation Based on the Assessment of the Example 5.2

Based on the given directed graph (Figure 5.2), we are examining the probability of knowing the concept C_j at skill level Lk. There is a set of responses $R_i = \{Q_1, Q_2, Q_3 \ldots. Q_n\}$, i = 1, 2, 3..,n. Each element of R_i is the response to the question that verifies knowledge about one or more concepts. There is a correct response denoted by $Q_{r,}$ and incorrect response denoted by $\overline{Q}_r$, r is an integer number that indicates index of the questions. Based on the given directed graph (Figure 5.2), find $P(C_j^{Lk}|R_i)$, Lk =1, 2,3,4,5,6 indicates the skill level of the concept C_j. In the directed graph (Figure 5.2), there are two types of nodes, questions and concepts, and directed links between them. There is a label

Lk on the directed link from the question node q_r to the concept node C_j. The label Lk means to answer q_i correctly, the concept C_j must be known at skill level b. The concept C_j could be source node C_s or target node C_t, which they are implied by the directed link that connects them. The label Lk on the directed link from the source concept C_s to the target concept C_t means to know C_t at skill level Lk, C_s must be known at skill level 2.

Considering that:

- $P(Q_r|\overline{C}_j^{Lk}) = g$, when there is dependency between Q_r and C_j
- $P(\overline{Q}_r|C_j^{Lk}) = m$
- $P(C_j^{Lk}) = d$

a) find $P(C_j^{Lk}|R_i)$ for R_1, R_2, R_3

$R_1 = Q_1Q_2Q_3\ Q_4$.

$R_2 = \overline{Q}_1\overline{Q}_2\overline{Q}_3\overline{Q}_4$.

$R_3 = \overline{Q}_1Q_2Q_3\overline{Q}_4$.

The solution

From the given information:

- $P(C_j^{Lk}) = d$
- $P(\overline{C}_j^{Lk}) = 1 - P(C_j^{Lk}) = 1 - d$
- Since $P(Q_r|\overline{C}_j^{Lk}) = g$. "$g$" is a value of error as same as "e" related to the asked question q_r, then $P(Q_r|C_j^{Lk}) = 1 - e = 1 - g$
- Since, in a special case where $P(\overline{Q}_r|C_j^{Lk}) = m$, "$m$" is the error value related to the asked question q_r, then $P(\overline{Q}_r|\overline{C}_j^{Lk}) = 1 - e = 1 - m$.

"g" presents the lucky guess, "m" presents mistake and "e" is the error. We assumed "e" is referring to any kind of errors such as the lucky guess or the mistake or the type of question which it could be direct question type or indirectly question asked about the concept skill at certain skill level C_j^{Lk}).

By Using Bayes' Theorem

$$P(C_j^{Lk}|R_i) = \frac{P\left(R_i \middle| C_j^{Lk}\right) * P\left(C_j^{Lk}\right)}{P\left(R_i \middle| C_j^{Lk}\right) P\left(C_j^{Lk}\right) + P\left(R_i \middle| \bar{C}_j^{Lk}\right) P\left(\bar{C}_j^{Lk}\right)}$$

- C_j^{Lk} denotes knowing the concept C_j number j at skill level Lk

- $P(C_j^{Lk})$ is the un conditional probability of knowing the concept C_j at skill level k, it is the initial probability of knowing the concept C_j^{Lk}. It is just the rate of the correct responses to the questions asked about the concept C_j^{Lk}.

- R_i is the set of the responses to the questions asked about the concept C_j^{Lk} .

- $P(R_i|C_j^{Lk})$ is the probability the responses (evidences) on the condition of knowing the concept C_j^{Lk}.

- $P\left(R_i \middle| \bar{C}_j^{Lk}\right)$ is the probability of the responses at skill level k, on the condition of not knowing the concept C_j .

- $P\left(\bar{C}_j^{Lk}\right)$ is the unconditional probability of not knowing the concept C_j at skill level k. It is the initial probability of not 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}.

1) $P(C_1^{L6}|R_1)$

   $R_1 = Q_1Q_2Q_3Q_4$

Based on the structure of the relation between the questions and the concepts, the questions q_2 , q_3 , q_4 are not related with C_1^{L6}

$P(C_1^{L6}|R_1) = P(C_1^{L6}|Q_1)$

$= P(Q_1|C_1^{L6}) = 1 - g$

2) $P(C_1^{L6}|R_2)$

$R_2 = \overline{Q}_1\overline{Q}_2\overline{Q}_3\overline{Q}_4$

Based on the structure of the relation between the questions and the concepts, the questions q_2 , q_3, q_4 are not related with C_1^{L6}

$P(C_1^{L6}|R_2) = P(C_1^{L6}|\overline{Q}_1) = P(\overline{Q}_1|C_1^{L6}) = m$

3) $R_3 = \overline{Q}_1 Q_2 Q_3 \overline{Q}_4$

Based on the structure of the relation between the questions and the concepts, the questions q_2 , q_3 , q_4 are not related with C_1^{L6}

$P(C_1^{L6}|R_3) = P(C_1^{L6}|\overline{Q}_1) = P(\overline{Q}_1|C_1^{L6}) = m$

4) $P(C_2^{L2}|R_1)$

$R_1 = Q_1Q_2Q_3\, Q_4$

- $P(C_2^{L2}|R_1) = P(C_2^{L2}|Q_2)$

  $= P(Q_2|C_2^{L2}) = 1 - g$

- Based on the structure of the relation between the questions and the concepts, q_1 , q_3 , q_4 are not related with C_2^{L2}

5) $P(C_2^{L2}|R_2)$

$R_2 = \overline{Q}_1\overline{Q}_2\overline{Q}_3\overline{Q}_4$

- Based on the structure of the relation between the questions and the concepts, Q_1, Q_3, Q_4 are not related with C_2^{L2}
- $P(C_2^{L2}|R_2) = P(C_2^{L2}|\overline{Q}_2) = P(\overline{Q}_2|C_2^{L2}) = m$

6) $P(C_2^{L2}|R_3)$

$R_3 = \overline{Q}_1 Q_2 Q_3 \overline{Q}_4$

Based on the structure of the relation between the questions and the concepts, the questions q_1, q_3, q_4 are not related with C_2^{L2}

$P(C_2^{L2}|R_3) = P(C_2^{L2}|Q_2) = P(Q_2 \middle| C_2^{L2}) = 1 - g$

7) $P(C_3^{L3}|R_1)$

$R_1 = Q_1 Q_2 Q_3 Q_4$

- Based on the structure of the relation between the questions and the concepts, the questions q_1, q_2, q_4 are not related with C_3^{L3}

$P(C_3^{L3}|R_1) = P(C_3^{L3}|Q_3) = P\left(Q_3 \middle| C_3^{L3}\right) = 1 - g$

8) $P(C_3^{L3}|R_2)$

$R_2 = \overline{Q}_1 \overline{Q}_2 \overline{Q}_3 \overline{Q}_4$

- Based on the structure of the relation between the questions and the concepts, the questions q_1, q_2, q_4 are not related with C_3^{L3}

$P(C_3^{L3}|R_2) = P(C_3^{L3}|\overline{Q}_3)$

$= P\left(\overline{Q}_3 \middle| C_3^{L3}\right) = m$

9) $P(C_3^{L3}|R_3)$

$R_3 = \overline{Q}_1 Q_2 Q_3 \overline{Q}_4$

- Based on the structure of the relation between the questions and the concepts, the questions q_1 , q_2 , q_4 are not related with C_3^{L3}

$P(C_3^{L3}|R_3) = P(C_3^{L3}|Q_3) = P(Q_3|C_3^{L3}) = 1 - g$

10) $P(C_4^{L4}|R_1)$

$R_1 = Q_1Q_2Q_3Q_4$

- Based on the structure of the relation between the questions and the concepts, the questions q_1 , q_2 , q_3 are not related with C_4^{L4}

$P(C_4^{L4}|R_1) = P(C_4^{L4}|Q_4) = P(Q_4|C_4^{L4}) = 1 - g$

11) $P(C_4^{L4}|R_2)$

$R_2 = \overline{Q}_1\overline{Q}_2\overline{Q}_3\overline{Q}_4$

$P(C_4^{L4}|R_2) = P(C_4^{L4}|\overline{Q}_4)$

$= P(\overline{Q}_4|C_4^{L4}) = m$

- $P(C_4^{L4}|R_3)$

$R_3 = \overline{Q}_1Q_2Q_3\overline{Q}_4$

$P(C_4^{L4}|R_3) = P(C_4^{L4}|\overline{Q}_4)$

$= P(\overline{Q}_4|C_4^{L4}) = m$

- Based on the structure of the relation between the questions and the concepts, the questions q_1 , q_2 , q_3 are not related with C_4^{L4}

12) $P(C_5^{L2}|R_1)$

$R_1 = Q_1Q_2Q_3\ Q_4$

- In the case of concept C_5^{L2}, which supports two concepts, one is the directly tested concept C_2^{L2}, and another one is the indirectly tested concept C_6^{L3}. In the computation of this example, we will consider only the directly tested concept, which is the concept C_2^{L2}, intuitively estimated, it is given the same value of C_2^{L2} multiplied by $P(C_5^{L2})$

$P(C_5^{L2}|R_1)$

$= P(C_2^{L2}|R_1)*P(C_5^{L2})$

$= P(C_2^{L2}|R_1)*\ d_{C_5^{L2}}$

- ($d_{C_5^{L2}}$ is the value of the unconditional probability of knowing the concept C_5^{L2})

$=P(C_2^{L2}|Q_2) *\ d_{C_5^{L2}}$

$= (1-g) *\ d_{C_5^{L2}}$

$=\ d_{C_5^{L2}} - d_{C_5^{L2}} * g$

13) $P(C_5^{L2}|R_2)$

$R_2 = \overline{Q}_1\overline{Q}_2\overline{Q}_3\overline{Q}_4$

- In the case of concept C_5^{L2}, which supports two concepts, one is the directly tested concept C_2^{L2}, and another one is the indirectly tested concept C_6^{L3}. In the computation of this example, we will consider only the directly tested concept, which is the concept C_2^{L2}, intuitively estimated, it is given the same value of C_2^{L2} multiplied by $P(C_5^{L2})$

$P(C_5^{L2}|R_2)$

$= P(C_2^{L2}|R_2)*P(C_5^{L2})$

$= P(C_2^{L2}|\overline{Q}_2) * \ d_{C_5^{L2}}$

$= P(\overline{Q}_2|C_2^{L2}) * d_{C_5^{L2}})$

($d_{C_5^{L2}}$ is the value of the unconditional probability of knowing the concept C_5^{L2})

$= m * d_{C_5^{L2}}$

14) $P(C_5^{L2}|R_3)$

$R_3 = \overline{Q}_1 Q_2 Q_3 \overline{Q}_4$

- In the case of concept C_5^{L2}, which supports two concepts, one is the directly tested concept C_2^{L2}, and another one is the indirectly tested concept C_6^{L3}. In the computation of this example we will consider only the directly tested concept, which is the concept C_2^{L2}, estimated by intuitive, it is given the same value of C_2^{L2} multiplied by $P(C_5^{L2})$

$P(C_5^{L2}|R_3) = P(C_2^{L2}|R_3) * P(C_5^{L2})$

$= P(C_2^{L2}|Q_2) * P(C_5^{L2})$

$= P(Q_2|C_2^{L2}) * d$

$= (1 - g) * \ d_{C_5^{L2}}$

$= \ d_{C_5^{L2}} - d_{C_5^{L2}} * g$

- ($d_{C_5^{L2}}$ is the value of the unconditional probability of knowing the concept C_5^{L2})

15) $P(C_6^{L2}|R_1)$

- In the case of concept C_6^{L2}, which supports two concepts, one is the directly tested concept C_3^{L3}, and another one is the indirectly tested concept C_7^{L2}. In the computation of this example, we will consider only the directly tested concept, which is the concept C_3^{L3}, intuitively estimated, it is given the same value of C_3^{L3} multiplied by $P(C_6^{L2})$.

$P(C_6^{L2}|R_1)$

$= P(C_3^{L3}|R_1)*P(C_6^{L2})$

$=P(C_3^{L3}|Q_3)*P(C_6^{L2})$

$= P(Q_3|C_3^{L3})*d$

$= 1 - g \; * d_{C_6^{L2}}$

$= d_{C_6^{L2}} - d_{C_6^{L2}} * g$

- ($d_{C_6^{L2}}$ is the value of the unconditional probability of knowing the concept C_6^{L2}

16) $P(C_6^{L2}|R_2)$

$P(C_6^{L2}|R_2) = P(C_3^{3}|R_2)*P(C_6^{L2})$

$= P(C_3^{L3}|R_2) * d$

$= P(C_3^{L3}|\overline{Q}_3) * d$

$= P(\overline{Q}_3|C_3^{L3}) * d$

$= m * \; d_{C_6^{L2}}$

- ($d_{C_6^{L2}}$ is the value of the unconditional probability of knowing the concept C_6^{L2})

17) $P(C_6^{L2}|R_3)$

$P(C_6^{L3}|R_3)$

$= P(R_3|C_3^{L3})*P(C_6^{L2})$

$= P(Q_3|C_3^{L3})*d$

$= (1-g) * \; d_{C_6^{L3}}$

- ($d_{C_6^{L3}}$ is the value of the unconditional probability of knowing the concept C_6^{L3})

18) $P(C_6^{L5}|R_1)$

- In the case of concept C_6^{L5}, which is supported at level 5 by only one concept C_5^{L2}, where the concept C_5^{L2} supports two concepts, one is the directly tested concept C_2^{L2}, and another one is the indirectly tested concept C_6^{L3}. In the computation of this example we will consider only the directly tested concept, which is the concept C_2^{L2}, intuitively estimated, it is given the same value of C_2^{L2} multiplied by $P(C_6^{L5})$.

$P(C_6^{L5}|R_1)$

$= P(C_5^{L2}|R_1)*P(C_6^{L5})$

$= P(C_2^{L2}|R_1)*P(C_6^{L5})$

$= P(C_2^{L2}|Q_2)*P(C_6^{L5})$

$= P(Q_2|C_2^{L2})*\ d_{C_6^{L5}}$

$= (1-g) * d_{C_6^{L5}}$

- ($d_{C_6^{L5}}$ is the value of the unconditional probability of knowing the concept C_6^{L5})

19) $P(C_6^{L5}|R_2)$

$P(C_6^{L5}|R_2)$

$= P(C_5^{L2}|R_2)*P(C_6^{L5})$

$= P(C_2^{L2}|R_2)*P(C_6^{L5})$

$= P(R_1|C_2^{L2})*\ d_{C_6^{L5}}$

$= P(\overline{Q}_2|C_2^{L2})*\ d_{C_6^{L5}}$

$= m * \mathrm{d}_{C_6^{L5}}$

- ($\mathrm{d}_{C_6^{L5}}$ is the value of the unconditional probability of knowing the concept C_6^{L5})

20) $P(C_6^{L5}|R_3)$

$= P(C_6^{L5}|R_3)$

$= P(C_6^{L5}|R_3)$

$= P(C_5^{L2}|R_3) * P(C_6^{L5})$

$= P(C_2^{L2}|R_3) * P(C_6^{L5})$

$= P(C_2^{L2}|Q_2) * \mathrm{d}_{C_6^{L5}}$

$= P(Q_2|C_2^{L2}) * \mathrm{d}_{C_6^{L5}}$

$= (1-g) * \mathrm{d}_{C_6^{L5}}$

- ($\mathrm{d}_{C_6^{L5}}$ is the value of the unconditional probability of knowing the concept C_6^{L5})

21) $P(C_7^{L2}|R_1)$

- In the case of concept C_7^{L2}, which supports only one directly tested concept C_3^{L3}, intuitively estimated, it is given the same value of C_3^{L3} multiplied by $P(C_7^{L2})$

$= P(C_3^{L3}|R_1) * P(C_7^{L2})$

$= P(C_3^{L3}|Q_3) * \mathrm{d}_{C_7^{L2}}$

$= P(Q_3|C_3^{L3}) * \mathrm{d}_{C_7^{L2}}$

- ($\mathrm{d}_{C_7^{L2}}$ is the value of the unconditional probability of knowing the concept C_7^{L2})

- $= (1-g)\mathrm{d}_{C_7^{L2}}$

22) $P(C_7^{L2}|R_2)$

- In the case of concept C_7^{L2}, which supports only one directly tested concept C_3^{L3}, intuitively estimated, it is given the same value of $P(C_3^{L3})$ multiplied by $P(C_7^{L2})$

$P(C_7^{L2}|R_2)$

$= P(C_3^{L3}|R_2)*P(C_7^{L2})$

$= P(C_3^{L3}|\overline{Q}_3)*\ d_{C_7^{L2}}$

$= P(\overline{Q}_3|C_3^{L3})*\ d_{C_7^{L2}}$

$= m*\ d_{C_7^{L2}}$

($d_{C_7^{L2}}$ is the value of the unconditional probability of knowing the concept C_7^{L2})

23) $P(C_7^{L2}|R_3)$

- In the case of concept C_7^{L2}, which supports only one directly tested concept C_3^{L3}, it is given the same value of C_3^{L3} multiplied by $P(C_7^{L2})$

- $P(C_7^{L2}|R_3)$

$= P(C_3^{L3}|R_3)*P(C_7^{L2})$

$= P(C_3^{L3}|Q_3)*\ d_{C_7^{L2}}$

$= P(Q_3|C_3^{L3})*\ d_{C_7^{L2}}$

- ($d_{C_7^{L2}}$ is the value of the unconditional probability of knowing the concept C_7^{L2})

$= (1-g)*\ d_{C_7^{L2}}$

24) $P(C_7^{L4}|R_1)$

- In the case of concept C_7^{L4}, which is supported at skill level 4 by more than one concept C_4^{L2}, and C_6^{L2}

- We know the probability of the support concepts C_4^{L2} & C_6^{L2} by the evidences R_1.

- The concept C_4^{L4} , is directly tested by q_4. The concept C_6^{L2} supports two concepts: the directly tested concept C_3^{L3} and the indirectly tested concept C_5^{L2}. In the computation of this example, we will consider only the directly tested concepts C_4^{L4} and C_3^{L3}, and we will not use the repeated questions if there is a repetition.
- $P(C_7^{L4}|R_1)$
- $$= \frac{(R_1|C_7^{L4})*P(C_7^{L4})}{P(R_1|C_7^{L4})*P(C_7^{L4})+P(R_1|\bar{C}_7^{L4})*P(\bar{C}_7^{L4})}$$

$$= \frac{P(R_1|C_4^{L2})*P(R_1|C_6^{L2})*P(C_7^{L4})}{P(R_1|C_4^{L2})*P(R_1|C_6^{L2})*P(C_7^{L4})+(R_1|\bar{C}_4^{L2})*[P(R_1|\bar{C}_6^{L2})*P(\bar{C}_7^{L4})}$$

$$= \frac{P(R_1|C_4^{L2})*P(R_1|C_3^{L3})*P(C_7^{L4})}{P(R_1|C_4^{L2})*P(R_1|C_3^{L3})*P(C_7^{L4})+(R_1|\bar{C}_4^{L2})*P(R_1|\bar{C}_3^{L3})*P(\bar{C}_7^{L4})}$$

$$= \frac{P(Q_4|C_4^{L2})*P(Q_3|C_3^{L3})*P(C_7^{L4})}{P(Q_4|C_4^{L2})*P(Q_3|C_3^{L3})*P(C_7^{L4})+P(Q_4|\bar{C}_4^{L2})*P(Q_3|\bar{C}_3^{L3})*P(\bar{C}_7^{L4})}$$

$$= \frac{(1-g)*(1-g)P(C_7^{L4})}{[(1-g)*(1-g)P(C_7^{L4})]+[g*g*P(\bar{C}_7^{L4})]}$$

- $$= \frac{(1-2gd+g^2)d_{C_7^{L4}}}{(1-2gd \quad ^2)d_{C_7^{L4}}]+[g^2*\left(1-d_{C_7^{L4}}\right)}$$

- ($d_{C_7^{L4}}$ is the value of the unconditional probability of knowing the concept C_7^{L4})

25) $P(C_7^{L4}|R_2)$

- $P(C_7^{L4}|R_2)$

$$= \frac{(R_2|C_7^{L4})*P(C_7^{L4})}{P(R_2|C_7^{L4})*P(C_7^{L4})+P(R_2|\bar{C}_7^{L4})*P(\bar{C}_7^{L4})}$$

$$=\frac{P\left(R_2\middle|C_4^{L2}\right)*P\left(R_2\middle|C_6^{L2}\right)*P(C_7^{L4})}{P\left(R_2\middle|C_4^{L2}\right)*P\left(R_2\middle|C_6^{L2}\right)*P(C_7^{L4})+\left(R_2\middle|\bar{C}_4^{L2}\right)*P\left(R_2\middle|\bar{C}_6^{L2}\right)*P(\bar{C}_7^{L4})}$$

$$=\frac{P\left(R_2\middle|C_4^{L2}\right)*P\left(R_2\middle|C_3^{L3}\right)*P(C_7^{L4})}{P\left(R_2\middle|C_4^{L2}\right)*P\left(R_2\middle|C_3^{L3}\right)*P(C_7^{L4})+\left(R_2\middle|\bar{C}_4^{L2}\right)*P\left(R_2\middle|\bar{C}_3^{L3}\right)*P(\bar{C}_7^{L4})}$$

$$=\frac{P(\bar{Q}_4|C_4^{L2}) * P(\bar{Q}_3|C_3^{L3}) * P(C_7^{L4})}{P(\bar{Q}_4|C_4^{L2}) * [P(\bar{Q}_3|C_3^{L3}) * P(C_7^{L4}) + P(\bar{Q}_4|\bar{C}_4^{L2})] * [P(\bar{Q}_3|\bar{C}_3^{L3}) * P(\bar{C}_7^{L4})}$$

$$=\frac{m*m*P\left(C_7^{L4}\right)}{m*m*P\left(C_7^{L4}\right)+[(1-m)*(1-m)*P\left(\bar{C}_7^{L4}\right)]}$$

$$=\frac{m^2*d_{C_7^{L4}}}{(m^2*d_{C_7^{L4}})+(1-2m+m^2)*(1-d_{C_7^{L4}})}$$

$$=\frac{m^2 * d_{C_7^{L4}}}{(m^2 * d_{C_7^{L4}}) + (1 - 2m + m^2 - d_{C_7^{L4}} + 2md_{C_7^{L4}} - m^2 d_{C_7^{L4}})}$$

26) $P(C_7^{L4}|R_3)$

$$=\frac{P\left(R_3\middle|C_7^{L4}\right)*P(C_7^{L4})}{P\left(R_3\middle|C_7^{L4}\right)*P(C_7^{L4})+P\left(R_3\middle|\bar{C}_7^{L4}\right)*P(\bar{C}_7^{L4})}$$

$$=\frac{P\left(R_3\middle|C_4^{L2}\right)*P\left(R_3\middle|C_6^{L2}\right)*P(C_7^{L4})}{P\left(R_3\middle|C_4^{L2}\right)*P\left(R_3\middle|C_6^{L2}\right)*P(C_7^{L4})+\left(R_3\middle|\bar{C}_4^{L2}\right)*P\left(R_3\middle|\bar{C}_6^{L2}\right)*P(\bar{C}_7^{L4})}$$

$$=\frac{P\left(R_3\middle|C_4^{L2}\right)*[P\left(R_3\middle|C_3^{L3}\right)*P(C_7^{L4})}{P\left(R_3\middle|C_4^{L2}\right)]*[P\left(R_3\middle|C_3^{L3}\right)*P(C_7^{L4})+\left(R_3\middle|\bar{C}_4^{L2}\right)]*P\left(R_3\middle|\bar{C}_3^{L3}\right)*P(\bar{C}_7^{L4})}$$

$$=\frac{P(\bar{Q}_4|C_4^{L2}) * P(Q_3|C_3^{L3}) * P(C_7^{L4})}{P(\bar{Q}_4|C_4^{L2}) * P(Q_3|C_3^{L3}) * P(C_7^{L4}) + P(\bar{Q}_4|\bar{C}_4^{L2}) * P(Q_3|\bar{C}_3^{L3}) * P(\bar{C}_7^{L4})}$$

$$=\frac{m * (1 - g) * P(C_7^{L4})}{m * (1 - g) * P(C_7^{L4}) + [(1 - m) * g * P(\bar{C}_7^{L4})]}$$

$$=\frac{m * (1 - g) * d_{C_7^{L4}}}{m * (1 - g) * d_{C_7^{L4}} + [(1 - m) * g * \left(1 - d_{C_7^{L4}}\right)]}$$

- ($d_{C_7^{L4}}$ is the value of the unconditional probability of knowing the concept C_7^{L4})

27) $P(C_8^{L5}|R_1)$

- In the case of concept C_8^{L5}, which is supported at level 5 by more than one concept, C_5^{L2} and C_7^{L2}. The concept C_5^{L2} supports two concepts, one is the directly tested concept C_2^{L2}, and another one is the indirectly tested concept C_6^{L3}. The concept C_7^{L2} supports the concept C_3^{L3}, which is directly tested by q_3. In the computation of this example, we will consider only the directly tested concept, which is the concept C_2^{L2}, and the concept C_3^{L3}.
- Therefore, we know the probability of the support concepts C_5^{L2} & C_7^{L2} by the evidences R_1.

$P(C_8^{L5}|R_1)$

$$= \frac{P(R_1|C_8^{L5}) * P(C_8^{L5})}{P(R_1|C_8^{L5}) * P(C_8^{L5}) + P(R_1|\bar{C}_8^{L5}) * P(\bar{C}_8^{L5})}$$

$$= \frac{P(R_1|C_5^{L2}) * P(R_1|C_7^{L2}) * P(C_8^{L5})}{P(R_1|C_5^{L2}) * P(R_1|C_7^{L2}) * P(C_8^{L5}) + P(R_1|\bar{C}_5^{L2}) * [P(R_1|\bar{C}_7^{L2}) * P(\bar{C}_8^{L5})]}$$

$$= \frac{P(R_1|C_2^{L2}) * P(R_1|C_3^{L3}) * P(C_8^{L5})}{P(R_1|C_2^{L2}) * P(R_1|C_3^{L3}) * P(C_8^{L5}) + P(R_1|\bar{C}_2^{L2}) * P(R_1|\bar{C}_3^{L3}) * P(\bar{C}_8^{L5})}$$

$$= \frac{P(Q_2|C_2^{L2}) * P(Q_3|C_3^{L3}) * P(C_8^{L5})}{P(Q_2|C_2^{L2}) * P(Q_3|C_3^{L3}) * P(C_8^{L5}) + P(Q_2|\bar{C}_2^{L2}) * P(Q_3|\bar{C}_3^{L3}) * P(\bar{C}_8^{L5})}$$

$$= \frac{(1-g) * (1-g) * P(C_8^{L5})}{[(1-g) * (1-g) * P(C_8^{L5})] + [g * g * P(\bar{C}_8^{L5})]}$$

- ($d_{C_8^{L5}}$ is the value of the unconditional probability of knowing the concept C_8^{L5})

$$= \frac{(1-2gd+g^2)*\, d_{C_8^{L5}}}{(1-2gd+g^2)*\, d_{C_8^{L5}} + g^2*\left(1-d_{C_8^{L5}}\right)}$$

28) $P(C_8^{L5}|R_2)$

$$= \frac{P\left(R_2\middle|C_8^{L5}\right)*P(C_8^{L5})}{P\left(R_2\middle|C_8^{L5}\right)*P(C_8^{L5})+P\left(R_2\middle|\bar{C}_8^{L5}\right)*P(\bar{C}_8^{L5})}$$

$$= \frac{P\left(R_2\middle|C_5^{L2}\right)*P\left(R_2\middle|C_7^{L2}\right)*P(C_8^{L5})}{P\left(R_2\middle|C_5^{L2}\right)*P\left(R_2\middle|C_7^{L2}\right)*P(C_8^{L5})+P\left(R_2\middle|\bar{C}_5^{L2}\right)*\left[P\left(R_2\middle|\bar{C}_7^{L2}\right)*P(\bar{C}_8^{L5})\right]}$$

$$= \frac{P\left(R_2\middle|C_2^{L2}\right)*P\left(R_2\middle|C_3^{L3}\right)*P(C_8^{L5})}{P\left(R_2\middle|C_2^{L2}\right)*P\left(R_2\middle|C_3^{L3}\right)*P(C_8^{L5})+P\left(R_2\middle|\bar{C}_2^{L2}\right)*P\left(R_2\middle|\bar{C}_3^{L3}\right)*P(\bar{C}_8^{L5})}$$

$$= \frac{P\left(\bar{Q}_2\middle|C_2^{L2}\right) * P\left(\bar{Q}_3\middle|C_3^{L3}\right) * P\left(C_8^{L5}\right)}{\left[P\left(\bar{Q}_2\middle|C_2^{L2}\right) * P\left(\bar{Q}_3\middle|C_3^{L3}\right) * P\left(C_8^{L5}\right)\right] + \left[P\left(\bar{Q}_2\middle|\bar{C}_2^{L2}\right) * P\left(\bar{Q}_3\middle|\bar{C}_3^{L3}\right) * P\left(\bar{C}_8^{L5}\right)\right]}$$

$$= \frac{m*m*P\left(C_8^{L5}\right)}{m*m*P\left(C_8^{L5}\right)+\left[(1-m)*(1-m)*P\left(\bar{C}_8^{L5}\right)\right]}$$

$$= \frac{m^2*d_{C_7^{L4}}}{(m^2*d_{C_8^{L5}})+(1-2m+m^2)*(1-d_{C_8^{L5}})}$$

$$= \frac{m^2 * d_{C_7^{L4}}}{(m^2 * d_{C_8^{L5}}) + (1 - 2m + m^2 - d_{C_8^{L5}} + 2md_{C_7^{L4}} - m^2 d_{C_8^{L5}})}$$

29) $P(C_8^{L5}|R_3)$

$$= \frac{P\left(R_3\middle|C_8^{L5}\right)*P(C_8^{L5})}{P\left(R_3\middle|C_8^{L5}\right)*P(C_8^{L5})+P\left(R_3\middle|\bar{C}_8^{L5}\right)*P(\bar{C}_8^{L5})}$$

$$= \frac{P\left(R_3\middle|C_5^{L2}\right)*P\left(R_3\middle|C_7^{L2}\right)*P(C_8^{L5})}{P\left(R_3\middle|C_5^{L2}\right)*P\left(R_1\middle|C_7^{L2}\right)*P(C_8^{L5})+P\left(R_3\middle|\bar{C}_5^{L2}\right)*P\left(R_3\middle|\bar{C}_7^{L2}\right)*P(\bar{C}_8^{L5})}$$

$$= \frac{P\left(R_3\middle|C_2^{L2}\right)*P\left(R_3\middle|C_3^{L3}\right)*P(C_8^{L5})}{P\left(R_3\middle|C_2^{L2}\right)*P\left(R_3\middle|C_3^{L3}\right)*P(C_8^{L5})+P\left(R_3\middle|\bar{C}_2^{L2}\right)*P\left(R_3\middle|\bar{C}_3^{L3}\right)*P(\bar{C}_8^{L5})}$$

$$= \frac{P\left(Q_2\middle|C_2^{L2}\right) * P\left(Q_3\middle|C_3^{L3}\right) * P\left(C_8^{L5}\right)}{P\left(Q_2\middle|C_2^{L2}\right) * P\left(Q_3\middle|C_3^{L3}\right) * P\left(C_8^{L5}\right) + P\left(Q_2\middle|\bar{C}_2^{L2}\right) * P\left(Q_3\middle|\bar{C}_3^{L3}\right) * P\left(\bar{C}_8^{L5}\right)}$$

$$= \frac{(1-g)*(1-g)*P\left(C_8^{L5}\right)}{[(1-g)*(1-g)*P\left(C_8^{L5}\right)]+[g*g*P\left(\bar{C}_8^{L5}\right)]}$$

- ($d_{C_8^{L5}}$ is the value of the unconditional probability of knowing the concept C_8^{L5})

$$= \frac{(1-2gd\quad ^2)*\, d_{C_8^{L5}}}{(1-2gd+g^2)*\, d_{C_8^{L5}} + g^2*\left(1-d_{C_8^{L5}}\right)}$$

b) Show the effecting of changing the value d on the behavior of $P(C_j^{Lk}|R_i)$, consider the values:

- $P(Q_r|\bar{C}_j^{Lk}) = g = 0.2$, when there is dependency between Q_r and C_j, "r" indicated the index of the question.
- $P(\bar{Q}_r|C_j^{Lk}) = m = 0.2$
- $P(C_j^{Lk}) = d$
- d = 0, d = 0.25, d = 0.50, d = 0.75, d = 1
- $P(\bar{C}_j^{Lk}) = 1 - d$

The solution: The effecting of changing the value of d on the behavior of $P(C_j^{Lk}|R_i)$ is shown in Figures 5.3, 5.4 and 5.5.

$P(C_j^{Lk}|R_i)$ for R = 1, 2, 3 is illustrated in Figure 5.3, Figure 5.4 and Figure 5.5 respectively. The figures show behavior of the probabilities of knowing the concepts C_1^6, C_2^2, C_3^3, C_4^4, C_5^2 C_6^2 , C_6^5, C_7^2 , C_7^4 and C_8^5. The first bundle at the left side in Figure 5.3, Figure 5.4 and Figure 5.8 show the effect of the unconditional probability value $P(\ C_J^{LK}) = 0$.

The last bundle at the right side in the chart shows the effect of the unconditional probability value P(C_J^{LK}) $= 1$.

The third bundle in the middle (left to right) side shows the effect of the unconditional probability value P(C_J^{LK}) $= 0.5$.

Figure 5.3 The Probabilities of the Concepts $P(C_j^{Lk} | R_1)$

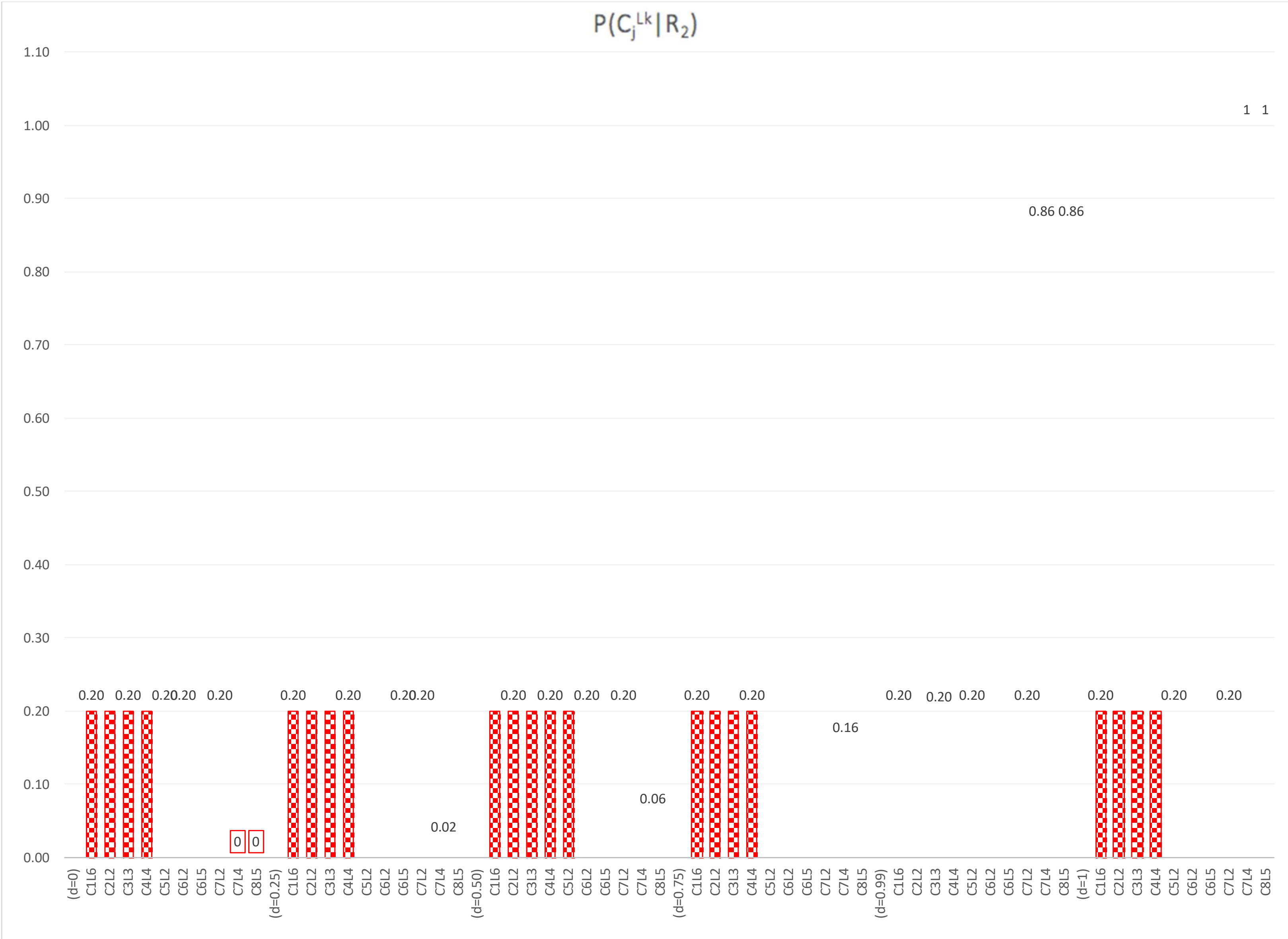


Figure 5.4 The Probabilities of the Concepts $P(C_j^{Lk}|\ R_2)$

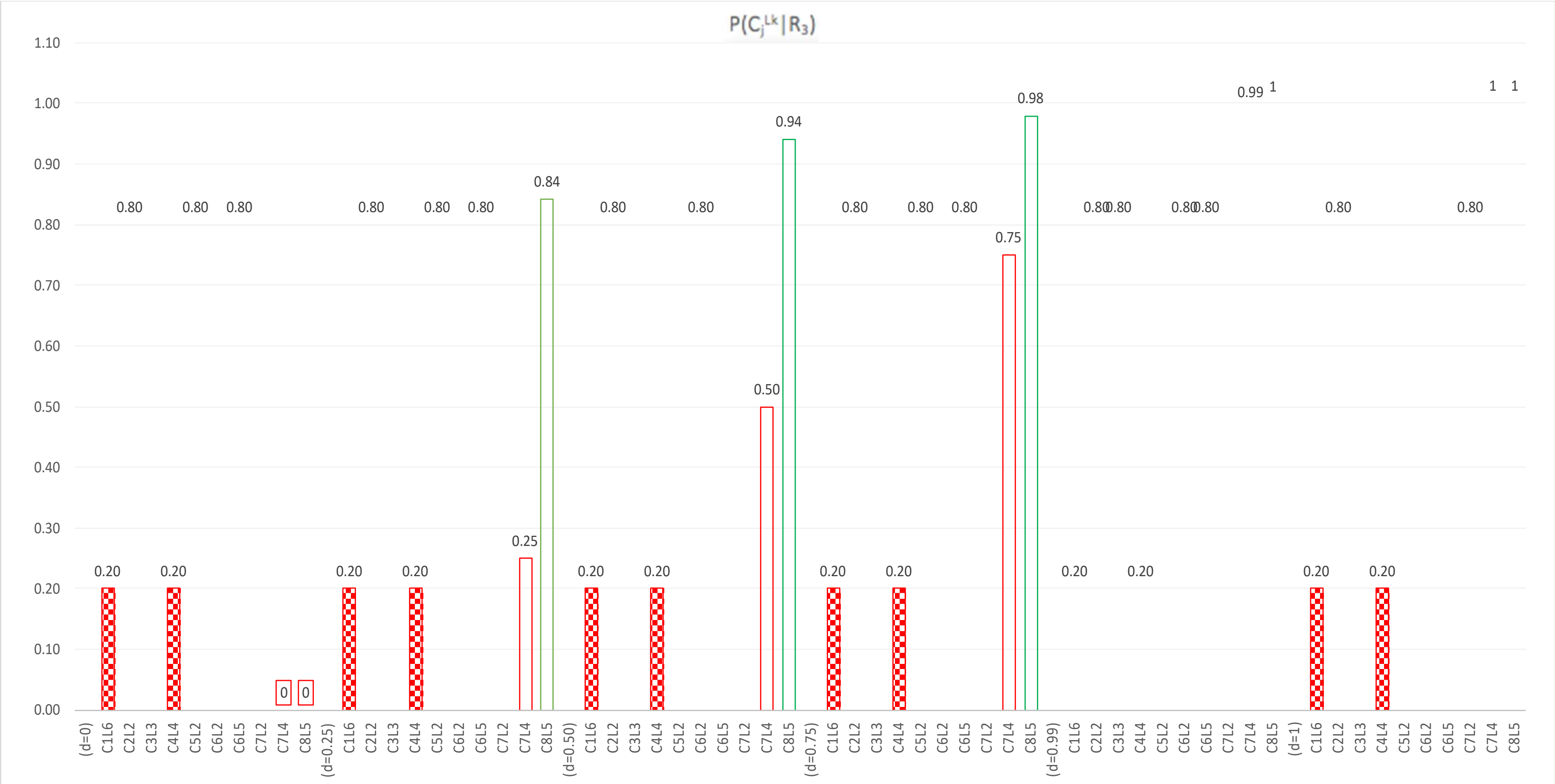


Figure 5.5 The Probabilities of the Concepts $P(C_j^{Lk}|\ R_3)$

5.5 Illustration of the calculation of the probability of knowing the concepts at skill levels based on the proposed methods

A directed Graph G (such as Figure 5.6) is given providing the dependency between a set of questions and a set of concepts. In G, there are two types of nodes: the question nodes denoted by q_i, and the concept nodes denoted by C_j. When there is a directed link Lk labelled from the question node q_r to the concept node C_j (indicated by solid arrow), it means that the ability to answer q_i correctly is dependent on knowing the concept C_j correctly at skill level k. When there is a directed link (indicated by line arrow) with label Lk from the source concept C_s to the target concept C_t, it means that knowing C_t correctly at skill level Lk is dependent on knowing C_s correctly at skill level 2. (We omit labeling the source end, as we imply that it is always level 2). Graph G shows these dependency relations.

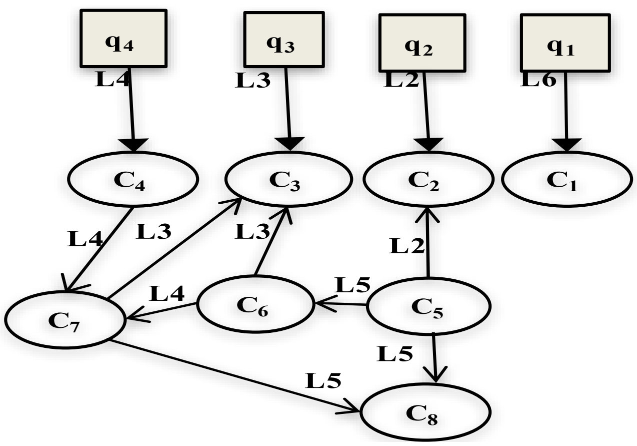


Figure 5.6 The Assessment Structure

As an input, we also have a set of responses denoted by an evidence vector $R_i = \{Q_1, Q_2,.. \overline{Q}_r,\dots Q_n\}$. Each element of R_i is the response to a question that verifies knowledge about of one or more concepts. A correct response is denoted by Q_r, and an incorrect response is denoted by $\overline{Q}_r$, "r" is an integer number that indicates the index of the question. We are now examining the probability of knowing various concepts given an evidence vector. Our formal question is:

Given the dependency Graph G, and specific response vector R_i (as evidence finds the conditional probability of knowing the concept C_j at level Lk that is $P(C_j^{Lk}|R_i)$, where Lk =1, 2,3,4,5,6.

Some additional probabilities are also given. Consider that:

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

$P(\overline{Q}_r|C_j^{Lk})$ =m, when there is a dependency between Q_r and C_j

$P(C_j^{Lk})$ = d.

We can find $P(C_j^{Lk}|R_i)$ for three different response vectors R_1, R_2, R_3, etc. In this example, I illustrate the estimation method with a specific example- how to find $P(C_8^{L5}|R_{3,3})$ for the specific response vectors $R_{3,3} = \{\overline{Q}_1, Q_2, Q_3\}$ for the graph given in Figure 5.6.

The solution

From the given information:

- $P(C_j^{Lk}) = d$
- $P(\bar{C}_j^{Lk}) = 1 - P(C_j^{Lk}) = 1 - d$
- Since $P(Q_r|\bar{C}_j^{Lk}) = g$. "$g$" is an error portion, which is "e". Thus, g = e. "g" is a value of error as same as "e" related to the asked question q_r, then $P(Q_r|C_j^{Lk}) = 1 - e = 1 - g$.
- Thus, the probability of the correct answer given that knowing the concept is $P(Q_r|C_j^{Lk}) = 1 - g$.
- Since $P(\bar{Q}_r|C_j^{Lk}) = m$. "$m$" is an error portion, which is "e". Thus, m = e.
- Since in a special case where $P(\bar{Q}_r|C_j^{Lk}) = m$. "$m$" is the error value related to the asked question q_r, then $P(\bar{Q}_r|\bar{C}_j^{Lk}) = 1 - P(Q_r|\bar{C}_j^{Lk}) = 1 - g = 1 - e = 1 - m$
- "g" presents the lucky guess, "m" presents a mistake and "e" is the error. We assumed "e" refers to any kind of error, such as the lucky guess, the mistake, or the type of question (direct question type or indirect question) asked about the concept at a certain skill level C_j^{Lk}.
- $R_{3,3} = \bar{Q}_1, Q_2, Q_3$

Using Bayes Theorem

$$P(C_8^{L5}|R_{3,3}) = \frac{P\left(R_{3,3}\middle|C_8^{L5}\right) * P(C_8^{L5})}{P\left(R_{3,3}\middle|C_8^{L5}\right) * P(C_8^{L5}) + P\left(R_{3,3}\middle|\bar{C}_8^{L5}\right) * P(\bar{C}_8^{L5})}$$

By the structure of the related concepts and the questions in the assessment, the directly tested concepts are:

The concept C_1^{L6} which is evaluated by only one question q_1.

The concept C_2^{L2} which is evaluated by only one question q_2 .

The concept C_3^{L3} which is evaluated by only one question q_3 .

The concept C_4^{L2} which is evaluated by only one question q_4 .

The question q_4 is not related to the concept C_8^{L5}.

To calculate the probability of knowing the concept C_8^{L5}, we should use the questions asked about each directly tested concept, for example,

$P\left(R_{3,3} \middle| C_1^{L6}\right) = P\left(\overline{Q}_1 \middle| C_1^{L6}\right) = P\left(C_1^{L6} \middle| \overline{Q}_1\right) = m.$

$P\left(\overline{Q}_1 \middle| C_1^{L6}\right) = m$ is given in the explanation of the question. The question q_1 is directly asked about the concept C_1^{L6}. In my proposal and by intuition, I proposed that, in this case, there is no previous information about the probability of knowing the concept and, if we test the concept at the first time by the q_1, then the probability of knowing the concept by given an evidence (response) $P\left(C_1^{L6} \middle| \overline{Q}_1\right)$ is equal to the probability of the response by given knowing the concept $P\left(\overline{Q}_1 \middle| C_1^{L6}\right)$, which is equal to m.

If we use Bayes' Theorem from one evidence, it will affect the behavior of this $P(C_8^{L5}|R_{3,3})$ incorrectly. If we have previous information about the probability of knowing the concept, then we can use Bayes' Theorem. However, it depends on the purpose of the exam and the evaluator.

$P\left(R_{3,3} \middle| C_2^{L2}\right) = P(Q_2 | C_2^{L2}) = P(C_2^{L2} | Q_2) = 1 - g$

$P(R_{3,3}|C_3^{L3}) = P(Q_3|C_3^{L3}) = P(C_3^{L2}|Q_3) = 1 - g$

The proof and justification is given later.

Now, let us calculate the probability of knowing the concept C_8^{L5}, which could be evaluated indirectly. Let's start by using the probability of knowing the indirectly evaluated concept C_8^{L5}, given the set of all events (responses) $R_{3,3}$ to arrive the exact related questions, which ask about the supported (related) directly tested concepts. The set of the responses $R_{3,3}$ will be replaced by the set of the responses to the exact questions asked about the related concepts to the concept C_8^{L5}.

$P(C_8^{L5}|R_{3,3})$

$$= \frac{P(R_{3,3}|C_8^{L5}) * P(C_8^{L5})}{P(R_{3,3})}$$

$$= \frac{P(R_{3,3}|C_8^{L5}) * P(C_8^{L5})}{P(R_{3,3}|C_8^{L5}) * P(C_8^{L5}) + P(R_{3,3}|\bar{C}_8^{L5}) * P(\bar{C}_8^{L5})}$$

The concept C_8^{L5}, is supported at level 5 by more than one concepts C_5^{L2} and C_7^{L2}.

$$= \frac{P(R_{3,3}|C_5^{L2}) * P(R_{3,3}|C_7^{L2}) * P(C_8^{L5})}{P(R_{3,3}|C_5^{L2}) * P(R_{3,3}|C_7^{L2}) * P(C_8^{L5}) + P(R_{3,3}|\bar{C}_5^{L2}) * P(R_{3,3}|\bar{C}_7^{L2}) * P(\bar{C}_8^{L5})}$$

The concept C_5^{L2} supports two concepts, one is the directly tested concept C_2^{L2}, and another one is the indirectly tested concept C_6^{L3}. The concept C_6^{L3} supports two concepts, one is the directly tested concept C_3^{L3}, and another one is the indirectly tested concept C_7^{L2}. The concept C_7^{L2} supports one directly tested concept C_3^{L3}. The concept C_3^{L3} is directly tested by the question q_3, and, because C_3^{L3} is supported by two concepts that refer to the concerned concept C_8^{L5} by three concepts, the question q_3 will be repeated independently 3 times in the computation of $P(C_8^{L5}|R_{3,3})$. Thus, in the derivation of Bayes' Theorem in this solution, we consider all the related concepts to the concept C_8^{L5} even though they are not directly tested by the corresponded question or they are not directly related to the directly tested concepts. Also, we consider the repeated question, which asks for many concepts, even with the short and long paths to the question that indirectly tests the concept C_8^{L5}. Therefore, the computation will be as the following:

$P(C_8^{L5}|R_{3,3})$

$$= \frac{P\left(R_{3,3}\middle|C_8^{L5}\right) * P(C_8^{L5})}{P(R_{3,3})}$$

$$= \frac{P\left(R_{3,3}\middle|C_8^{L5}\right) * P(C_8^{L5})}{P\left(R_{3,3}\middle|C_8^{L5}\right) * P(C_8^{L5}) + P\left(R_{3,3}\middle|\bar{C}_8^{L5}\right) * P(\bar{C}_8^{L5})}$$

(The concept C_8^{L5}, is supported at level 5 by more than one concept C_5^{L2} and C_7^{L2}).

$$= \frac{P\left(R_{3,3}\middle|C_5^{L2}\right) * P\left(R_{3,3}\middle|C_7^{L2}\right) * P(C_8^{L5})}{P\left(R_{3,3}\middle|C_5^{L2}\right) * P\left(R_{3,3}\middle|C_7^{L2}\right) * P(C_8^{L5}) + P\left(R_{3,3}\middle|\bar{C}_5^{L2}\right) * P\left(R_{3,3}\middle|\bar{C}_7^{L2}\right) * P(\bar{C}_8^{L5})}$$

"The concept C_5^{L2} supports the directly tested concept C_2^{L2} and the indirectly tested the concept C_6^{L2}. The concept C_7^{L2}, supports the directly tested concept C_3^{L3}).

$$= \frac{P\left(R_{3,3}\middle|C_2^{L2}, C_6^{L3}\right) * P\left(R_{3,3}\middle|C_3^{L3}\right) * P\left(C_8^{L5}\right)}{P\left(R_{3,3}\middle|C_2^{L2}, C_6^{L3}\right) * P\left(R_{3,3}\middle|C_3^{L3}\right) * P\left(C_8^{L5}\right) + P\left(R_{3,3}\middle|\bar{C}_2^{L2}, \bar{C}_6^{L3}\right) * P\left(R_{3,3}\middle|\bar{C}_3^{L3}\right) * P\left(\bar{C}_8^{L5}\right)}$$

$$= \frac{P\left(R_{3,3}\middle|C_2^{L2}, C_6^{L3}\right) * P\left(R_{3,3}\middle|C_3^{L3}\right) * P\left(C_8^{L5}\right)}{P\left(R_{3,3}\middle|C_2^{L2}, C_6^{L3}\right) * P\left(R_{3,3}\middle|C_3^{L3}\right) * P\left(C_8^{L5}\right) + P\left(R_{3,3}\middle|\bar{C}_2^{L2}, \bar{C}_6^{L3}\right) * P\left(R_{3,3}\middle|\bar{C}_3^{L3}\right) * P\left(\bar{C}_8^{L5}\right)}$$

(The indirectly tested the concept C_6^{L2} supports the indirectly tested concept C_7^{L2}, and the directly tested concept C_3^{L3})

$$= \frac{P(R_{3,3}|C_2^{L2}, C_3^{L3},\ C_7^{L2}) * P(R_{3,3}|C_3^{L3}) * P(C_8^{L5})}{P(R_{3,3}|C_2^{L2}, C_3^{L3}, C_7^{L2}) * P(R_{3,3}|C_3^{L3}) * P(C_8^{L5}) + P(R_{3,3}|\bar{C}_2^{L2}, \bar{C}_3^{L3}, \bar{C}_7^{L2}) * P(R_{3,3}|\bar{C}_3^{L3}) * P(\bar{C}_8^{L5})}$$

(The indirectly tested the concept C_7^{L2} supports the directly tested concept C_3^{L3})

$$= \frac{P(R_{3,3}|C_2^{L2}, C_3^{L3}, C_3^{L3}) * P(R_{3,3}|C_3^{L3}) * P(C_8^{L5})}{P(R_{3,3}|C_2^{L2}, C_3^{L3}, C_3^{L3}) * P(R_{3,3}|C_3^{L3}) * P(C_8^{L5}) + P(R_{3,3}|\bar{C}_2^{L2}, \bar{C}_3^{L3}, \bar{C}_3^{L3}) * P(R_{3,3}|\bar{C}_3^{L3}) * P(\bar{C}_8^{L5})}$$

(The concept C_2^{L2}, is evaluated by the question q_2, and the concept C_3^{L3} is evaluated by the question q_3)

The derivation of the computation of the probability of knowing the concept C_8^{L5} is completed such that

$$= \frac{P(Q_2|C_2^{L2}) * P(Q_3|C_3^{L3}) * P(Q_3|C_3^{L3}) * P(Q|C_3^{L3}) * P(C_8^{L5})}{P(Q_2|C_2^{L2}) * P(Q_3|C_3^{L3}) * P(Q_3|C_3^{L3}) * P(Q_3|C_3^{L3}) * P(C_8^{L5}) * P(C_8^{L5}) + P(Q_2|\bar{C}_2^{L2}) * P(Q_3|\bar{C}_3^{L3}) * P(Q_3|\bar{C}_3^{L3}) * P(Q_3|\bar{C}_3^{L3}) * P(\bar{C}_8^{L5})}$$

$$= \frac{(1-g)*(1-g)*(1-g)*(1-g)*d}{[(1-g)*(1-g)*(1-g)*(1-g)*d]+[g*g*g*g*(1-d)]}$$

- The computation of the probability $P(C_8^{L5}|R_{3,3})$ if d = 0.25 is as follows:

$P(C_8^{L5}|R_{3,3})$

$$= \frac{(1-0.2)*(1-0.2)*(1-0.2)*(1-0.2)*d}{[(1-0.2)*(1-0.2)*(1-0.2)*(1-0.2)*d]+[0.2*0.2*0.2*0.2*(1-d)]}$$

$$= \frac{0.8*0.8*0.8*0.8*d}{[0.8*0.8*0.8*0.8*d]+[0.2*0.2*0.2*0.2*(1-d)]}$$

$$= \frac{0.8*0.8*0.8*0.8*0.25}{[0.8*0.8*0.8*0.8*0.25]+[0.2*0.2*0.2*0.2*(0.75)]}$$

$$= \frac{0.1}{[0.1]+[0.2*0.2*0.2*0.2*0.75]} = 0.001 = 0$$

- The computation of the probability of knowing the concept $P(C_8^{L5}|R_{3,3})$ if d = 0.75 is as follows:

$P(C_8^{L5}|R_{3,3})$ if d = 0.75

$P(C_8^{L5}|R_{3,3})$

$$= \frac{0.8*0.8*0.8*0.8*d}{[0.8*0.8*0.8*0.8*d]+[0.2*0.2*0.2*0.2*(1-d)]}$$

$$= \frac{0.8 * 0.8 * 0.8 * 0.8 * 0.75}{[0.8 * 0.8 * 0.8 * 0.8 * 0.75] + [0.2 * 0.2 * 0.2 * 0.2 * 0.25]}$$

$$= \frac{0.3}{0.3+0.0004} = 1$$

Let's find the computation of the probability of knowing the concept $P(C_8^{L5}|R_{3,3})$, where d is the Optimal Value.

d = The Optimal Value = The ratio of the number of correct responses to the number of questions refer to the concerned concept C_8^{L5} =

$$\frac{\text{Total number of the correct responses to the debendent questions}}{\text{Total number of the dependent questions}} = 4/4 = 1$$

$P(C_8^{L5}|R_{3,3})$

$$= \frac{0.8*0.8*0.8*0.8*d}{[0.8*0.8*0.8*0.8*d]+[0.2*0.2*0.2*0.2*(1-d)]}$$

$$= \frac{0.8 * 0.8 * 0.8 * 0.8 * 1}{[0.8 * 0.8 * 0.8 * 0.8 * 1] + [0.2 * 0.2 * 0.2 * 0.2 * 0]}$$

$$= \frac{0.4}{0.4+0} = 1$$

5.5.1 Example 5.3: The Value of the Unconditional Probability of Knowing the Concept $P(C_j^{Lk})$

This example introduces the case of the evaluated concept by many evidences with different unconditional probability values $P(C_j^{Lk})$. The investigation is to observe the effect of the value of $P(C_j^{Lk})$ on the behavior of $P(C_j^{Lk}|R_{3,i})$ to show the optimal unconditional probability. $R_{3,i}$ is the set of the responses to the questions i = 1, 2, 3..,n. Lk = 1,2,3,4,5,6 indicates the skill of the concept C_j^{Lk}. We will study the effect of the 6 curves of the value $P(C_j^{Lk})$ on the behavior of $P(C_j^{Lk}|R_{3,i})$ by calculating $P(C_j^{Lk}|R_{3,i})$ of the responses successively.

The question is:

Show the effecting of the different values of $P(C_{\mathbf{8}}^{L5})$ on the behavior of $P(C_{\mathbf{8}}^{L5}|R_{3,i})$ to prove the optimal value of d.

$R_{3,1} = \{\overline{Q}_1\}$

$R_{3,2} = \{\overline{Q}_1, Q_2\}$

$R_{3,3} = \{\overline{Q}_1, Q_2, Q_3\}$

$R_{3,4} = \{\overline{Q}_1, Q_2, Q_3, \overline{Q}_4\}$

$R_{3,5} = \{\overline{Q}_1, Q_2, \overline{Q}_3\}$

$R_{3,6} = \{\overline{Q}_1, \overline{Q}_2, \overline{Q}_3, \overline{Q}_4\}$

The values of the variables of $P(C_{\mathbf{8}}^{L5})$ = d as the following:

1. Curve a1:

g = m = 0.2

d_1= 0

2. Curve a2:

g = m = 0.2

d_2 =0.25

3. Curve a3:

g = m = 0.2

d_3 = 0.5

4. Curve a4:

g = m = 0.2

d_4 = 0.75

5. Curve a5:

m = g = m = 0.2

d_5 = 0.99

6. Curve a6:

m = g = m = 0.2

d_6 = 0.99

7. Curve a7:

m = g = m = 0.2

d_7 = Optimal value of d, which is the unconditional probability of knowing the concept.

a) Find $P(C_8^{L5}|R_{3,i})$ for the following evaluation of the sequence of the learner responses.

$R_{3,1} = \{\overline{Q}_1\}$

$R_{3,2} = \{\overline{Q}_1, Q_2\}$

$R_{3,3} = \{\overline{Q}_1, Q_2, Q_3\}$

$R_{3,4} = \{\overline{Q}_1, Q_2, Q_3, \overline{Q}_4\}$

$R_{3,5} = \{\overline{Q}_1, Q_2, \overline{Q}_3\}$

$R_{3,6} = \{\overline{Q}_1, \overline{Q}_2, \overline{Q}_3, \overline{Q}_4\}$

The solution:

To show the optimal value of d, we tried to avoid the repeated questions. For this reason, in the derivation of the computation of the probability of knowing the concept C_8^{L5}, we assume there is no repeating in the questions that indirectly asked about the concept C_8^{L5}. Therefore, if there is more than one concept related to the concerned (evaluated) concept and to the directly tested concept of the same question, we consider the question only once.

As proposed earlier, $P(C_{\mathbf{8}}^{L5})$ is the unconditional probability of knowing the concept C_j^{Lk}, which is the initial probability of knowing the concept $C_{\mathbf{8}}^{L5}$. I suggest that it is just the ratio of the correct responses to the questions asked about the concept C_j^{Lk}.

$$P(C_j^{Lk}) = \frac{|R_Q|}{|Q_q|}$$

In the case that there is no previous information about the probability of knowing the concept $C_{\mathbf{j}}^{LK}$, and the concept $C_{\mathbf{j}}^{LK}$has never been tested, then it is given the unconditional probability of equally likely knowing and not knowing the concept $C_{\mathbf{j}}^{LK}$.

If there is only one assessment with many questions, then the most appropriate value is the ratio of the correct responses to the questions asked about the concept C_j^{Lk}.

Figure 5.7 shows the effect of the different values of $P(C_{\mathbf{8}}^{L5})$ on the behavior of $P(C_{\mathbf{8}}^{L5}|R_{3,i})$

As observed if $P(C_{\mathbf{8}}^{L5}) = 0$ or 1, the effecting of the true evidences R will be cancelled.

If the unconditional probability $P(C_{\mathbf{8}}^{L5}) = 0$, then the numerator, which is the multiplication of the evidences (responses) on the condition of knowing the concept C_j^{Lk}, $P(R_{3,i}|C_{\mathbf{8}}^{L5})$ will be canceled with the unconditional probability $P(C_{\mathbf{8}}^{L5})$. Therefore, the conditional probabilities $P(R_{3,i}|C_{\mathbf{8}}^{L5})$ will be zero, and the result of $P(C_{\mathbf{8}}^{L5})\ |R_{3,i})$ will be equal to zero even though there are correct responses. For example, in the illustration in Figure 5.7, the case of $P(C_{\mathbf{8}}^{L5}) = d_1 = 0$ and $R_{3,5} = \{\overline{Q}_1, Q_2, \overline{Q}_3\}$, The evaluation result $P(C_{\mathbf{8}}^{L5}|R_{3,5}) = 0$ even though the number of correct responses is equal to (or even more than) the number of the incorrect responses.

$P(C_{\mathbf{8}}^{L5}|R_{3,5})$

$$= \frac{P(R_{3,i})\ |C_j^{Lk}) * P(C_j^{Lk})}{P(R_{3,i})\ |C_j^{Lk}) * P(C_j^{Lk}) + P(R_{3,i})\ |\overline{C}_j^{Lk}) * P(\overline{C}_j^{Lk})}$$

$$= \frac{P(Q_2|C_j^{Lk}) * P(\overline{Q}_3|C_j^{Lk}) * d_1}{P(Q_2|C_j^{Lk}) * P(\overline{Q}_3|C_j^{Lk}) * d_1 + P(Q_1|\overline{C}_j^{Lk}) * P(\overline{Q}_3|\overline{C}_j^{Lk}) * (1-d_1)}$$

$$= \frac{0.8 * 0.2 * 0}{0.8 * 0.2 * 0 + 0.2 * 0.8 * 1}$$

$= 0$

Likewise, If the unconditional probability $P(C_j^{Lk}) = d_6 = 1$, then the right side in the denominator, which is the side of the evidences (responses) on the condition of not knowing the concept $C_j^{Lk} = 0$, because they will be canceled with the unconditional probability of not knowing the concept $P(\bar{C}_j^{Lk}) = 1 - d = 1 - 1 = 0$. Therefore, the final result of $P(C_{\mathbf{8}}^{L5}|R_{3,6})$ will be equal to 1, even though the most are the incorrect responses.

For example: in the chart of $R_{3,6}$ in the case of $P(C_{\mathbf{8}}^{L5}) = d_6 = 1$ and $R_{3,6} = \{\bar{Q}_1, \bar{Q}_2, \bar{Q}_3, \bar{Q}_4\}$, the $P(C_{\mathbf{8}}^{L5}|R_{3,6}) = 1$

$P(C_{\mathbf{8}}^{L5}|R_{3,6})$

$=$

$$\frac{P(R_{3,6}|C_{\mathbf{8}}^{L5}) * P(C_{\mathbf{8}}^{L5})}{P(R_{3,6}|C_{\mathbf{8}}^{L5}) * P(C_{\mathbf{8}}^{L5}) + P(R_{3,6}|\bar{C}_{8}^{L5}) * P(\bar{C}_{8}^{L5})}$$

$$= \frac{P(\bar{Q}_2|C_{\mathbf{8}}^{L5}) * P(\bar{Q}_3|C_{\mathbf{8}}^{L5}) * d_6}{P(\bar{Q}_2|C_{\mathbf{8}}^{L5}) * P(\bar{Q}_3|C_{\mathbf{8}}^{L5}) * d_6 + P(\bar{Q}_2|\bar{C}_{8}^{L5}) * P(\bar{Q}_3|\bar{C}_{8}^{L5}) * (1 - d_6)}$$

$$= \frac{0.8 * 0.2 * 0.2 * 0.2 * 1}{0.8 * 0.2 * 0.2 * 0.2 * 1 + 0.2 * 0.8 * 0.8 * 0.8 * 0}$$

$= 1$

On the other hand, any value of $d \in (0.5, 1)$ changes the behavior of the $P(C_j^{Lk})$ correctly, but the problem is that the probability value tends to be close to the probability of knowing the concept, even though the not knowing evidences/responses are more than number of correct responses. The reason is that the absolute value of $P(C_j^{Lk})$ could be higher than the value of the one evidence of not knowing the concept. In other words, it tends to be close the probability of knowing the concept. Also, when $d \in (0, 0.5)$ it changes the behavior of the $P(C_j^{Lk})$ correctly, but the problem is that the probability value tends be close to the probability of not knowing the concept, even though the knowing evidences are higher. The reason is that the absolute value of $P(C_j^{Lk})$ could be less than the value of the probability of one evidence of knowing the concept. In other words, it tends to be close to the probability of not knowing the concept. For this reason, the value of the unconditional probability d should be optimal choice. If there is a previous assessment, the value d is replaced by the probability of knowing the concept by the learner, which was estimated from the previous assessment. In other words, if there is previous information about the probability of knowing the concept by the learner, then the d of the new assessment will be replaced with the previous probability of knowing the concept to find out the accurate evaluation of the probability of knowing the concept according to the new assessment. If we don't have any previous information, then the optimal d should be chosen. The optimal value is d = The ratio of the correct answers to the total questions asked about the concept. The effect of the optimal value is shown in the Figure 5.8. As evidence, it gives the correct value of $P(C_j^{Lk}|R_i)$.

The Justification of $\mathbf{P\left(Q_r \middle| C_j^{Lk}\right) = P\left(C_j^{Lk} \middle| Q_r\right)}$:

If we don't have any previous information about the probability of knowing the concept and we asked one question presented to the learner, then the conditional probability of knowing the concept will be equal to the conditional probability of the given response.

The proof of this fact could be shown as the following steps:

1. From the conditional probability

$$P\left(C_j^{Lk} \middle| Q_r\right) = \frac{P\left(C_8^{L5}\right) \cap P(Q_r)}{P(Q_r)}$$

2. $P\left(C_8^{L5}\right) \cap P(Q_r) = P\left(C_j^{Lk} \middle| Q_r\right) * P(Q_r)$

$P(Q_r)$ presents the correct response for the q_r, so $P(Q_r) = 1$

3. Thus, $P\left(C_8^{L5}\right) \cap P(Q_r) = P\left(C_j^{Lk} \middle| Q_r\right)$

Also,

4. $P\left(Q_r \middle| C_j^{Lk}\right) = \frac{P(Q_r) \cap P\left(C_8^{L5}\right)}{P(C_j^{Lk})}$
5. $P\left(C_8^{L5}\right) \cap P(Q_r) = P\left(Q_r \middle| C_j^{Lk}\right) * P(C_j^{Lk})$

$P\left(C_j^{Lk}\right)$ presents the probability for only one concept by using only one question, so

$P\left(C_j^{Lk}\right) = 1$

6. Thus, $P\left(C_8^{L5}\right) \cap P(Q_r) = P\left(C_j^{Lk} \middle| Q_r\right)$

From 3 & 6

7. $P\left(C_j^{Lk} \middle| Q_r\right) = P\left(Q_r \middle| C_j^{Lk}\right)$
8. By replacing the value in the formula, then

$P\left(C_j^{Lk} \middle| Q_r\right) = P\left(Q_r \middle| C_j^{Lk}\right)$

Thus, we use Bayes' Theorem if we have many questions or many evidences referring to the probability of knowing and the concept, and the optimal d is the rate of the correct response based on the total number of the questions. The equation is

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

In the case of concept C_8^{L5}, which is supported at level 5 by more than one concept; C_5^{L2}, and C_7^{L2}

From the structure of the assessment, the related questions (indirect relations) to the concept C_8^{L5} are q_2, q_3. Thus $R_{3,1}$ will not affect the probability of the concept $P(C_8^{L5})$, where $R_{3,1} = \overline{Q}_1$

Thus,

$P(C_8^{L5}|R_{3,1}) = P(C_8^{L5}|\overline{Q}_1) = P(C_8^{L5}) = d$

If we start with the probability of one response to the question q_2

$P(C_8^{L5}|R_{3,2}) = P(C_8^{L5}|\overline{Q}_1 Q_2) = P(C_8^{L5}|Q_2) = P(Q_2|C_8^{L5}) = 0.8$

then when the second related question q_3 is asked about the concept C_8^{L5}, we compute the probability of knowing the concept C_8^{L5} given $R_{3,3} = \{Q_2, Q_3\}$

In this illustration of using Bayes' Theorem, I will use the value of $d_2 = 0.25$ in the comparison between $R_{3,2} = \{Q_2, Q_3\}$ & $R_{3,3} = \{\overline{Q}_1, Q_2, Q_3\}$. In the derivation of the computation of the probability of knowing the concept C_8^{L5} given $R_{3,i}$, we assume there is no repeating in the questions that indirectly asked about the concept C_8^{L5}. Therefore, if there is more than one concept related to the concerned (evaluated) concept and to the directly tested concept of the same question, we consider the question only once.

$P(C_8^{L5}|R_{3,3})$

$$= \frac{P(R_{3,3}|C_8^{L5}) * P(C_8^{L5})}{P(R_{3,3}|C_8^{L5}) * P(C_8^{L5}) + P(R_{3,3}|\overline{C}_8^{L5}) * P(\overline{C}_8^{L5})}$$

$$= \frac{P(\overline{Q}_1|C_8^{L5}) * P(Q_2|C_8^{L5}) * P(Q_3|C_8^{L5}) * P(C_8^{L5})}{P(\overline{Q}_1|C_8^{L5}) * P(Q_2|C_8^{L5}) * P(Q_3|C_8^{L5}) * P(C_8^{L5}) + P(\overline{Q}_1|\overline{C}_8^{L5}) * P(Q_2|\overline{C}_8^{L5}) * P(Q_3|\overline{C}_8^{L5}) * P(\overline{C}_8^{L5})}$$

(From the structure of the assessment, the related questions (indirect relations) are q_2 and q_3 , but q_1 is not related with the concept C_8^{L5} either by direct or indirect relation.

Therefore, q_1 doesn't affect the probability of knowing the concept $P(C_8^{L5})$ and it was removed from the equation)

$$= \frac{P(Q_2|C_8^{L5}) * P(Q_3|C_8^{L5}) * P(C_8^{L5})}{P(Q_2|C_8^{L5}) * P(Q_3|C_8^{L5}) * P(C_8^{L5}) + P(Q_2|\bar{C}_8^{L5}) * P(Q_3|\bar{C}_8^{L5}) * P(\bar{C}_8^{L5})}$$

$$= \frac{P(Q_2|C_8^{L5}) * P(Q_3|C_8^{L5}) * P(C_8^{L5})}{P(Q_2|C_8^{L5}) * P(Q_3|C_8^{L5}) * P(C_8^{L5}) + P(Q_2|\bar{C}_8^{L5}) * P(Q_3|\bar{C}_8^{L5}) * P(\bar{C}_8^{L5})}$$

$$= \frac{0.8 * 0.8 * 0.25}{0.8 * 0.8 * 0.25 + 0.2 * 0.2 * 0.75}$$

= 0.842

As observed, the difference between $P(C_8^{L5}|R_{3,3})$ & $P(C_8^{L5}|R_{3,2})$ is $0.842 - 0.8 = 0.042$. It is small but the behavior of the $P(C_8^{L5}|R_{3,3})$ is increased because we have the second evidence (second correct response) of knowing the concept. Figure 5.7 shows the effecting of all the given values of d on $P(C_8^{Lk}|R_{3,i})$.

Now, I will illustrate using Bayes' Theorem from the first evidence (response). Starting using Bayes' Theorem from one evidence is not an appropriate solution. Let's try it:

$P(C_8^{L5}|R_{3,2})$

$$= \frac{P(R_{3,2}|C_8^{L5}) * P(C_8^{L5})}{P(R_{3,2}|C_8^{L5}) * P(C_8^{L5}) + P(R_{3,2}|\bar{C}_8^{L5}) * P(\bar{C}_8^{L5})}$$

(From the structure of the assessment, the related questions (indirect relations) are q_2 and q_3 , but q_1 is not related either by direct or indirect relation. Therefore, q_1 doesn't affect the probability of knowing the concept $P(C_8^{L5})$, and it was removed from the calculation).

$$= \frac{P(Q_2|C_8^{L5}) * P(C_8^{L5})}{P(Q_2|C_8^{L5}) * P(C_8^{L5}) + P(Q_2|\bar{C}_8^{L5}) * P(\bar{C}_8^{L5})}$$

$$= \frac{P\left(Q_2 \middle| C_8^{L5}\right) * d_2}{P\left(Q_2 \middle| C_8^{L5}\right) * d_2 + P\left(Q_2 \middle| \overline{C}_8^{L5}\right) * \overline{d}_2}$$

$$= \frac{0.8*0.25}{0.8*0.25+0.2*0.75}$$

$= 0.57$

Now let's use Bayes' Theorem to compute $P(C_8^{L5} | R_{3,3})$ by using the previous information.

$P(C_8^{L5} | R_{3,3})$

$$= \frac{P\left(R_{3,3} \middle| C_8^{L5}\right) * P(C_8^{L5})}{P\left(R_{3,3} \middle| C_8^{L5}\right) * P(C_8^{L5}) + P\left(R_{3,3} \middle| \overline{C}_8^{L5}\right) * P(\overline{C}_8^{L5})}$$

$$= \frac{P\left(\overline{Q}_1 \middle| C_8^{L5}\right) * P\left(Q_2 \middle| C_8^{L5}\right) * P\left(Q_3 \middle| C_8^{L5}\right) * P\left(C_8^{L5}\right)}{P\left(\overline{Q}_1 \middle| C_8^{L5}\right) * P\left(Q_2 \middle| C_8^{L5}\right) * P\left(Q_3 \middle| C_8^{L5}\right) * P\left(C_8^{L5}\right) + P\left(\overline{Q}_1 \middle| \overline{C}_8^{L5}\right) * P\left(Q_2 \middle| \overline{C}_8^{L5}\right) * P\left(Q_3 \middle| \overline{C}_8^{L5}\right) * P\left(\overline{C}_8^{L5}\right)}$$

(From the structure of the assessment, the related questions (indirect relations) are q_2 and q_3 , but q_1 is not related either by direct or indirect relation. Therefore, q_1 does not affect the probability of knowing the concept $P(C_8^{L5})$, and it was removed from the computation).

$$= \frac{P\left(Q_2 \middle| C_8^{L5}\right) * P\left(Q_3 \middle| C_8^{L5}\right) * P(C_8^{L5})}{P\left(Q_2 \middle| C_8^{L5}\right) * \left(Q_3 \middle| C_8^{L5}\right) * P(C_8^{L5}) + P\left(Q_2 \middle| \overline{C}_8^{L5}\right) * P\left(Q_3 \middle| \overline{C}_8^{L5}\right) * P(\overline{C}_8^{L5})}$$

$$= \frac{0.57*0.57*0.25}{0.57*0.57*0.25+0.43*0.43*0.75}$$

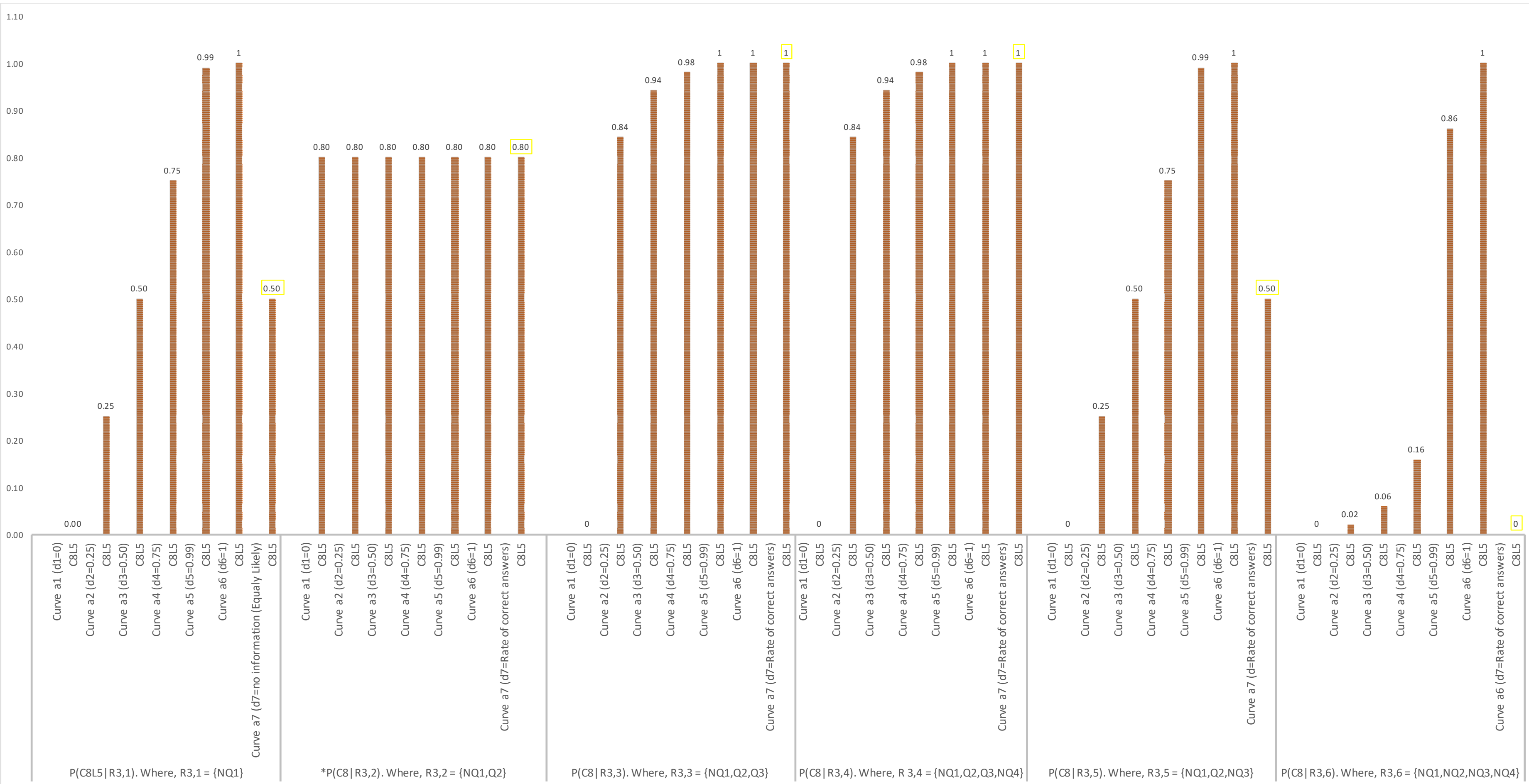
$$= \frac{0.081225}{0.081225+0.138675}$$

$$= \frac{0.081225}{0.2199} = 0.37$$

As observed, the probability of knowing the concept C_8^{L5} given by two evidences is decreased, even though the second evidence gave the second correct response, which indicates knowing the concept C_8^{L5}. The decrease of the probability of knowing the concept occurred because the value d of knowing the concept is 0.25, which leads to the value of the unconditional probability of not knowing the concept to be 0.75. This breaks the balance of the probabilities of knowing and unknowing the concept.

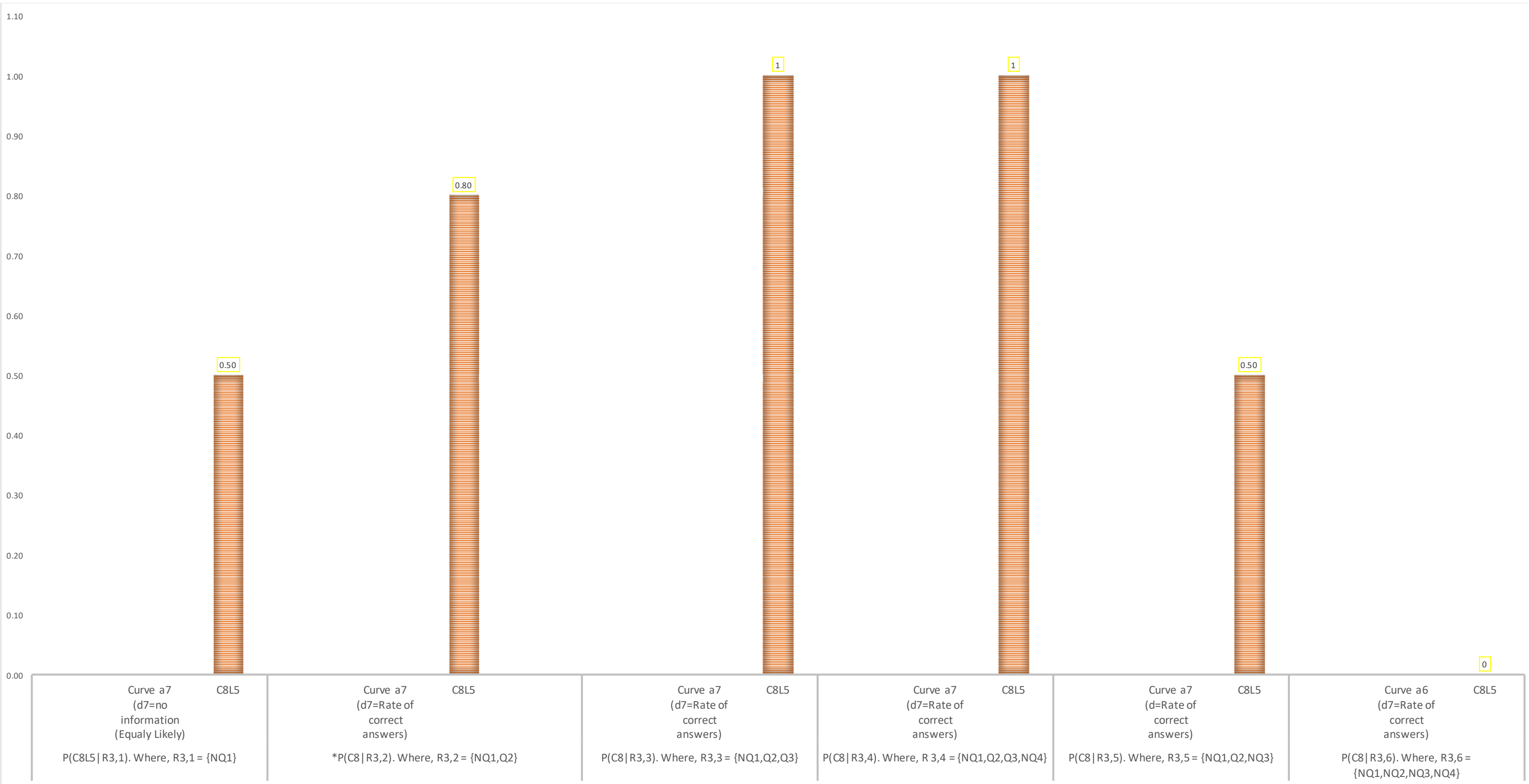
The effect of the probability of knowing the concept C_8^{L5} by starting using Bayes' Theorem from one question/evidence is shown in Figure 5.9.

As evidence on Figure 5.8, the optimal value of $P(C_8^{L5})$ is the ratio of the correct answers to the total questions asked about the concept $P(C_8^{L5})$. Figure 5.8 shows $P(C_8^{L5}|R_{3,i})$ by $P(C_8^{L5})$ is the optimal value, which is the ratio of the correct answers to the total questions asked about the concept. The first bar in the left shows the $P(C_8^{L5}|R_{3,i})$ when there is no information about the concept such that $P(C_8^{L5})$ is the equally likely of knowing and not knowing the concept.



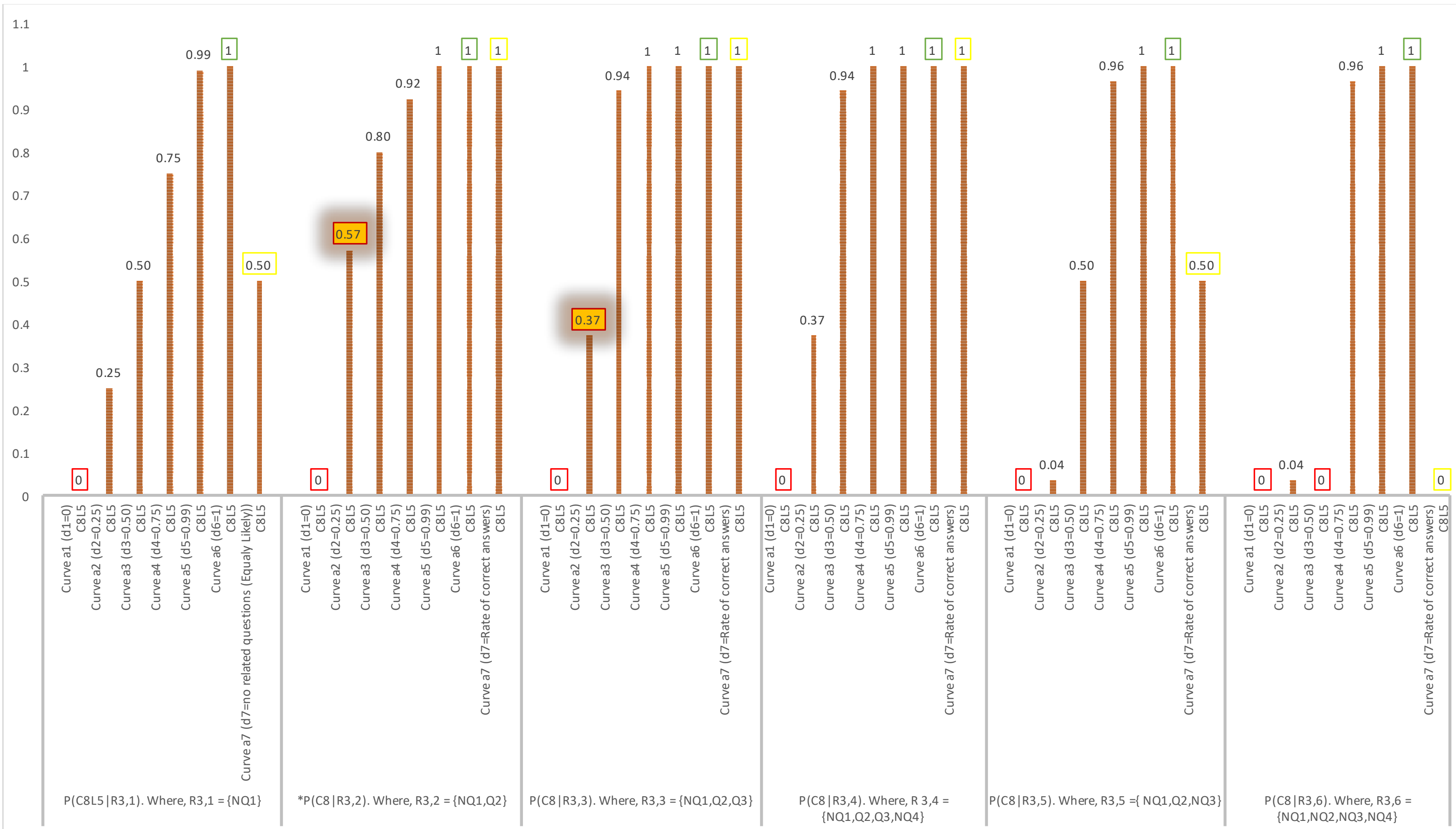
*He got correct response so the probability of knowing the concept by given 1 response is $P(C_8^{L5}|Q_2) = 1 - 0.2 = 0.8$

Figure 5.7 Probability of Knowing the Concept $P(C_8^{L5}|R_{3,i})$



* He got correct response so the probability of knowing the concept by given 1 response is $P(C_8^{L5}|Q_2) = 1 - 0.2 = 0.8$

Figure 5.8 The Effecting on the $P(C_8^{L5}|R_{3,i})$ by Using the $P(C_8^{L5})$ is the Optimal Value (Ratio of the Correct Answers to the Total Questions Asked About the Concept (If there is no Reference Questions d is the Equally Likely)



* He got correct response so the probability of knowing the concept by given 1 response is $P(C_8^{L5}|Q_2) = 1 - 0.2 = 0.8$

Figure 5.9 The Effecting on the $P(C_8^{L5})$ by Starting Using Bayes' Theorem from one question/evidence

5.5.2 Example 5.4: The Probability of Knowing a Concept Indirectly Evaluated

This example introduces the complex relation between the concepts in the assessment domain. The complexity is illustrated by the different cases of the existing of the concepts in a Concept Space.

Let the set of questions $Q_q = \{q_1, q_2\}$, asked about the set of concepts CS = $\{C_1^{LK}, C_2^{LK}, C_3^{LK}\}$ be such that question q_1 asks about concepts C_1^{LK}, C_2^{LK} and question q_2 asks about the concepts C_1^{LK}, C_2^{LK}, C_3^{LK}, Lk =1,2,3,4,5,6 indicates the skill of the concept C_j. Figure 5.10 illustrates this example:

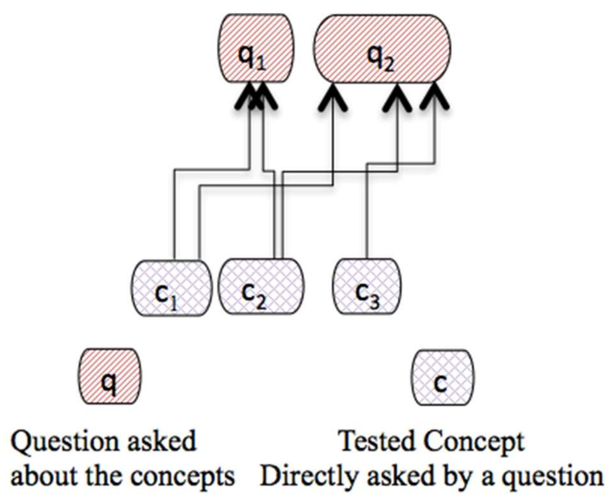


Figure 5.10 Set of Questions Asked About Set of Concepts

Suppose a learner answers the asked question, and the answer is correct. Then the probability of knowing the concept by given a correct response is $P(C_j^{LK}|Q_r) = (1-g_r)$, and the probability of not knowing the concept by given correct response is $P(\bar{C}_1^{LK}|Q_r) = g_r$. On the other hand, if the response to question q_r is incorrect, then the probability of knowing the concept C_j^{LK}, asked by the question q_r is $P(C_j^{LK}|\overline{Q}_r) = \mathcal{C}_r$, and the probability of not knowing the concept C_j^{LK} given by incorrect response is $P(\bar{C}_j^{LK}|\overline{Q}) = (1- \mathcal{C}_r)$. The two constants, $\mathcal{C}_r, g_r \in [0,1[$, are respectively called (careless) error probability and guessing probability at q_r. I suggest the error values such that, if the errors $\mathcal{C}_r, g_r$ are respectively given values 0.1 and 0.1, where r is an integer number that indicates index of the questions, we would find the probability of knowing the concepts according to the following cases:

1. Case 1

If the question q_1 is answered correctly, find the probability of knowing the concept C_1^{LK}. $P(C_1^{LK}|q_1)$. Figure 5.11 illustrates the case.

Solution:

Let $P(C_1^{LK})$ be the probability of knowing the concept C_1^{LK}.

Let Q_1 be a response to the question q_1 asked about concepts C_1^{LK}..

Given Q_1 is a correct answer, then, the probability of knowing the concept C_1^{LK}) given the correct response to question q_1 $P(C_1^{LK}|Q_1) = 1- 0.1= 0.9$

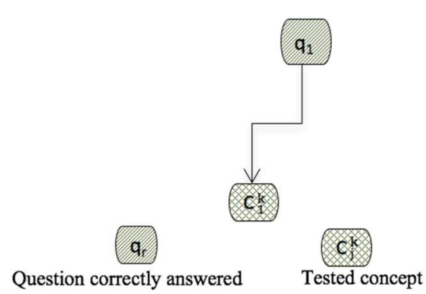


Figure 5.11 Example 5.4: Case 1: Correct Answer to the Question q_1

2. Case 2

If the question q_2 is answered correctly, find the probability of knowing the concept C_1^{LK}, $P(C_1^{LK})|Q_2)$. Figure 5.12 illustrates the case.

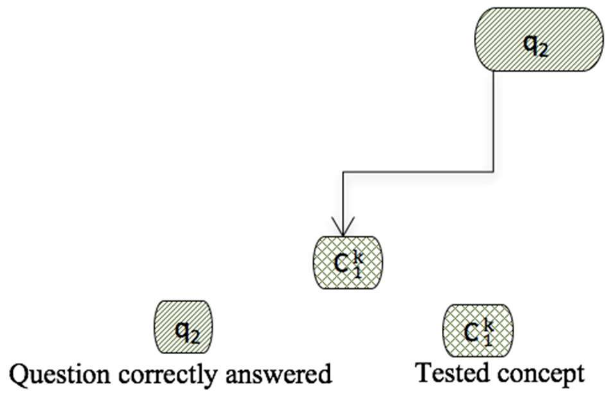


Figure 5.12 Example 5.4: Case 2: Correct Answer to the Question q_2

Solution:

Let $P(C_1^{LK})$ be the probability of knowing the concept C_1^{LK}.

Let Q_2 be a response to the question q_2 asked about a concept C_1^{LK}.

Given Q_2 is a correct answer and if the answer is correct, then the probability of knowing the concept C_1^{LK} is $P(C_1^{LK}|Q_r)= (1-g_r)$.

Therefore, the probability of knowing the concept C_1^{LK} given the correct response to question q_2, $p(C_1^{LK}|Q_2) = 1- 0.1= 0.9$.

3. Case 3

If the question q_2 is answered incorrectly, find the probability of knowing the concept C_1^{LK}, $P(C_1^{LK}|\overline{Q}_2)$. Figure 5.13 illustrates the case.

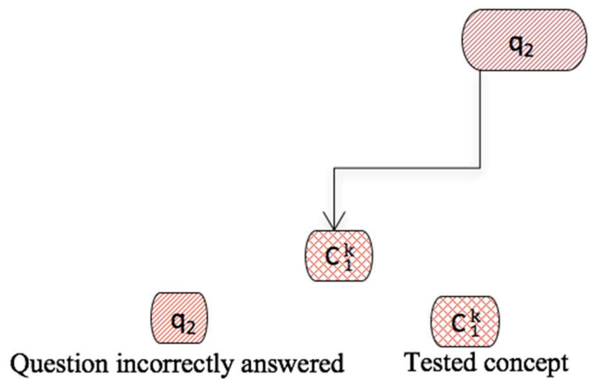


Figure 5.13 Example 5.4: Case 3: Wrong Answer to the Question q_2

- Let $P(C_1^{LK})$ be the probability of knowing the concept C_1^{LK}

- Let $\overline{Q}_2$ be a response to the question q_2 asked about the concepts C_1^{LK}

Given $\overline{Q}_2$ is an incorrect answer and if the answer is incorrect then the probability of knowing the concept $P(C_1^{LK}|\overline{Q}_2) = C_r = 0.1$, where C_r is the probability of careless error, r integer number referred to the question number r.

Therefore, the probability of knowing the concept C_1^{LK} given an incorrect response to a question q_2, $P(C_1^{LK}|\overline{Q}_2) = C_r = 0.1$.

4. Case 4

Given information data of the responses to the asked questions such that Q_1 is correct and $\overline{Q}_2$ is incorrect, find the probabilities $P(C_1^{LK}|Q_1, \overline{Q}_2)$, $P(C_2^{LK}|Q_1, \overline{Q}_2)$ and $P(C_3^{LK}|Q_1, \overline{Q}_2)$, Figure 5.14 illustrates the case.

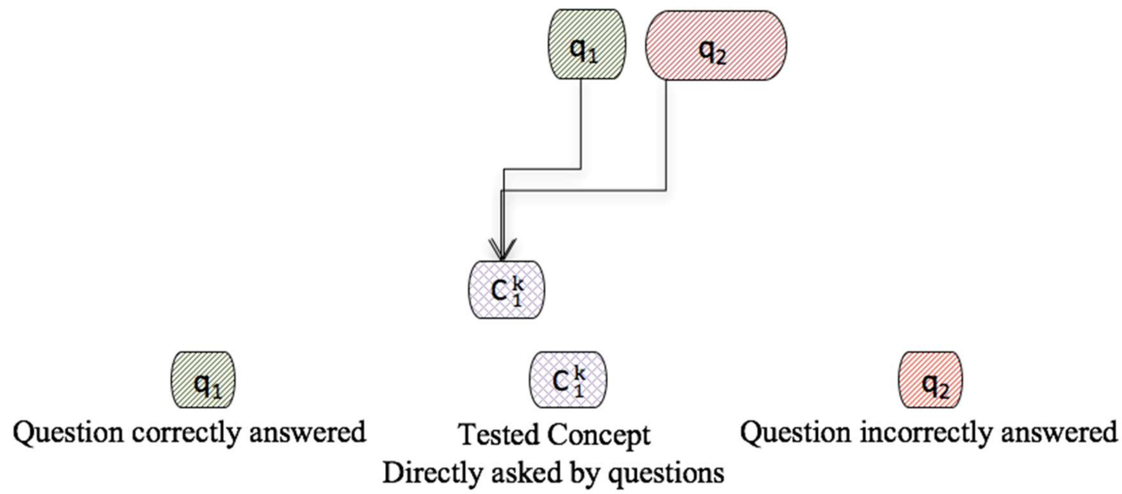


Figure 5.14 Example 5.4: Case 4: The Conflicted Answers

Solution:

Given:

- The responses data: Q_1 is correct and $\bar{Q}_2$ is incorrect.
- The probability of knowing the concept C_1^{LK} given response r_1 is $P(C_1^{LK}|Q_1) = 0.9$ (From the given problem statement).
- The probability of knowing the concept C_1^{LK} given response $\bar{Q}_2$ is $P(C_1^{LK}|\bar{Q}_2) = 0.1$ (From the given problem statement).
- The probability of not knowing the concept C_1^{LK} given response Q_1 is $P(\bar{C}_1^{LK}|Q_1) = 0.1$ (From the given problem statement).
- The probability of not knowing the concept $\bar{C}_1^{LK}$ given incorrect response $\bar{Q}_2$ is $P(\bar{C}_1^{LK}|\bar{Q}_2) = 0.9$ (From the given problem statement).

Let R be the set of events of the responses to the questions asked about a concept $\bar{C}_1^{LK}$.

$R = \{Q_1, \overline{Q}_2\}$

To calculate the probability of knowing the concept C_1^{LK} given the data of the responses R, we use Bayes' Theorem:

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \quad \rightarrow \quad \text{Basic Formula}$$

$$= \frac{P(B|A) * P(A)}{P(B|A) * P(A) + P(B|\overline{A}) * P(\overline{A})} \quad \rightarrow \quad \text{Extended Formula}$$

Replacing the notation we used in this dissertation then;

$$P(C_1^{LK}|Q_1, \overline{Q}_2) = \frac{P(Q_1, \overline{Q}_2|C_1^{LK}) * P(C_1^{LK})}{P(Q_1, \overline{Q}_2)}$$

$$= \frac{P(Q_1, \overline{Q}_2|C_1^{LK}) * P(C_1^{LK})}{P(Q_1, \overline{Q}_2|C_1) * P(C_1^{LK}) + P(Q_1, \overline{Q}_2|\overline{C}_1^{LK}) * P(\overline{C}_1^{LK})}$$

By replacing the notation of the two questions with one notation R as they are combined data.

$$P(C_1^{LK}|R) = \frac{P(R|C_1^{LK}) * P(C_1^{LK})}{P(R)}$$

$$= \frac{P(R|C_1^{LK}) * P(C_1^{LK})}{P(R|C_1^{LK}) * P(C_1^{LK}) + P(R|\overline{C}_1^{LK}) * P(\overline{C}_1^{LK})} ,$$

where

$$P(C_1^{LK}) = \frac{\text{Total number of the correct answers to the questions asked about the concept } C_1^{LK}}{\text{Total number of the questions asked about the concept } C_1^{LK}}$$

$$P(C_1^{LK}) = \frac{|R_Q|}{|Q_q|}$$

Thus,

$P(C_1^{LK}) = \frac{1}{2} = 0.5$ From Equation 3

$$P(\bar{C}_1^{LK}) = \frac{\text{Total number of the incorrect answers to the questions asked about the concept } C_1^{LK}}{\text{Total number of the questions asked about the concept } C_1^{LK}}$$

$P(\bar{C}_1^{LK}) = \frac{|R_{r\bar{q}}|}{|Q|}$ From Equation 4

$P(R|C_1^{LK})$ is the conditional probability of the event that the responses in R occur, conditional on the event of knowing the concept C_1^{LK}.

$P(R|\bar{C}_1^{LK})$ is the conditional probability of the event that the responses in R occur, conditional on the event of not knowing the concept C_1^{LK}.

Since Q_1 and $\bar{Q}_2$ are conditional unconditional events, wherein $Q_1 \in$ the set of correct answers and $\bar{Q}_2 \in$ the set of incorrect answers, then $P(R|C_1^{LK}) = P(Q_1) * P(\bar{Q}_2)$.

Thus, from the multiplication rules of independent events (The product rule) inferred from definition 3 in the book (Rohatgi & Ehsanes Saleh, 2015, p. 34).[14] Is:

$P(R\big|C_1^{LK}) = P(Q_1|C_1^{LK}) * P(\bar{Q}_2|C_1^{LK})$. By considering the condition of knowing the concept C_1^{LK}

$= 0.9 * 0.1 = 0.09$

$P(R|\bar{C}_1^{LK}) = P(Q_1|\bar{C}_1) * P(\bar{Q}_2|\bar{C}_1)$ By considering the condition of not knowing the concept C_1^{LK}

$P(R|\bar{C}_1^{LK})$) = 0.1 * 0.9 = 0.09

[14] http://www.wiley.com/WileyCDA/WileyTitle/productCd-111879964X.html

$P(R) = P(R|C_1^{LK}) * P(C_1^{LK}) + P(R|\bar{C}_1^{LK}) * P(\bar{C}_1^{LK})$ From the Theory Law of Total Probability

Thus,

$P(R) = 0.09*0.5 + 0.09*0.5 = 0.09$

By replacing the elements to Bayes' Theorem

$$P(C_1^{LK}|R) = \frac{P\left(R\middle|C_1^{LK}\right)*P(C_1^{LK})}{P(R)} \qquad \text{Basic Bayes' Formula}$$

$$= \frac{\left(R\middle|C_1^{LK}\right)*P(C_1^{LK})}{P\left(R\middle|C_1^{LK}\right)*P(C_1^{LK})+P\left(R\middle|\bar{C}_1^{LK}\right)*P(\bar{C}_1^{LK})} \qquad \text{Extended Bayes' Formula}$$

$$= \frac{0.09*0.5}{0.09} = 0.5$$

Similarly, the calculation for $P(C_2^{LK}|R)$ the probability of knowing the concept C_2^{LK}, has been evaluated based on the same two questions q_1 and q_2, where q_1 was answered correct and q_2 was answered wrong. These are denoted as Q_1 and $\bar{Q}_2$. The probability of knowing the concept C_2^{LK}, $P(C_2^{LK}) = 0.5$.

For C_3^{LK}, the probability of knowing the concept C_3^{LK}, $P(C_3^{LK}| \bar{Q}_2) = 0.1$, since it was asked by only q_2, which was answered wrong. Table 5.3 shows the computation probabilities $P(C_1^{LK})$, $P(C_2^{LK})$ and $P(C_3^{LK})$:

Table 5.3 Summary of the Probability Computation Result of Directly Tested Concepts - In Example 5.4 the Directly Tested Concepts Are C_1^{LK}, C_2^{LK}, C_3^{LK}

The concept	The related questions	The responses	The evaluation of responses	The probability of the concept by Bayes
C_1^{LK}	q_1, q_2	1,0	0.9, 0.1	0.5
C_2^{LK}	q_1, q_2	1,0	0.9, 0.1	0.5
C_3^{LK}	q_2	0	0.1	0.1

We conclude that the probability of knowing a concept C_1^{LK} that is evaluated based upon two questions and given conflicted responses shows that the probability of knowing the concept is equal to the probability of not knowing the concept. To determine the truth estimation, we give an error probability based on the type of the question. The estimation of the probability is increased if the question type is multiple choice. Also, the probability error will be assigned a higher value for the question indirectly asking about the target skill of the tested concept, rather than the question directly asking about the skill. In the experimental data, we calculate the probability of knowing the concept in the two cases: either, the probability of error is given the equal value in both direct question and indirect question, or the probability of the errors is given unequal value to the direct and indirect question. In this example, we finalized the calculation of all the concepts by using the same errors value 0.1, which were given in the question.

5. Case 5

Considering there are untested prerequisite sets of concepts, such that C_4^{LK} is a prerequisite of C_1^{LK}, and C_5^{LK} is a prerequisite of C_1^{LK} and C_2^{LK}, C_6^{LK} is a prerequisite for C_1^{LK}, C_2^{LK} and C_3^{LK}. C_7^{LK} is a prerequisite for C_3^{LK} and the responses to the questions R = $\{Q_1, \overline{Q}_2\}$ Find the probabilities $P(C_4^{LK}|Q_1,\overline{Q}_2)$, $P\left(C_5^{LK}|Q_1,\right)$ and $P(C_6^{LK}\ |Q_1,\overline{Q}_2)$, $P(C_7^{LK}|\overline{Q}_2)$. The result of the calculation of these three probabilities is illustrated in Table 5.3.

The calculation of the probabilities of knowing these concepts without considering the prerequisite relation will be in a simple calculation following the explanation of $P(C_1^{LK}|Q_1,\overline{Q}_2)$, $P(C_2^{LK}|Q_1,\overline{Q}_2)$ and $P(C_2^{LK}|\overline{Q}_2)$.

Thus,

$P(C_4^{LK}|Q_1,\overline{Q}_2) = 0.5$

$P\left(C_5^{LK}|Q_1,\overline{Q}_2\right) = 0.5$

$P(C_6^{LK}\middle|Q_1,\overline{Q}_2) = 0.5$

$P(C_7^{LK}\middle|\overline{Q}_2) \quad = 0.1$

Because these concepts are not directly tested and some of them are perhaps prerequisite to many tested concepts, such as the concepts C_5^{LK} and C_6^{LK} in the Figure 5.15, we should use the extended formula of Bayes' Theorem to calculate the probability of knowing the concept C_j^{LK}, which is an element in the set of DS. Also, if the concept C_j^{LK} is a prerequisite for a single concept, it should have a probability of knowing a concept based on the type of the supported tested concept whether it is in VKS or in VNS.

Thus, there are many cases to calculate the probability of knowing the concepts in DS. I study some cases as they are described in the following paragraph.

The existence of the prerequisite relationship between the directly tested concept and the indirectly tested concept implies that C_4^{LK}, C_5^{LK}, C_6^{LK}, and C_7^{LK}, are concepts in DS. Therefore, the suggested solution would be based upon the tested concept whether in VKS or in VNS. Figure 5.15 illustrates the concepts relation.

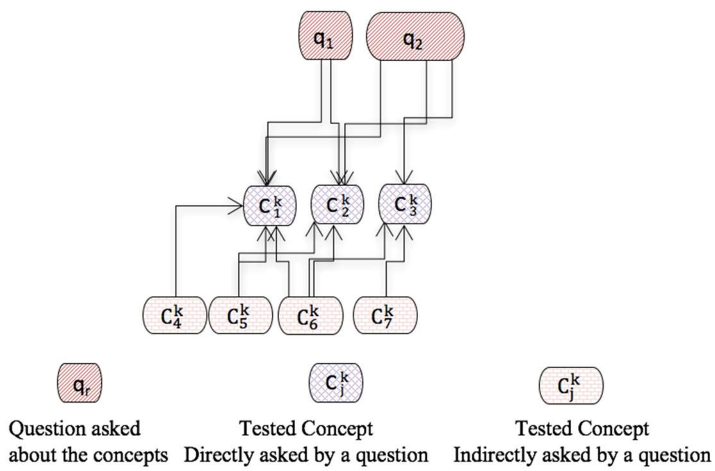


Figure 5.15 Example 5.4: Case 5: Considering the Data of Responses and The Related Concepts

Let the tested concept A be in the set of VS.

Let the related concept c be in the set DS of A.

1) If the related (supported) tested concept A is evaluated as a Verified Known Skill (VKS) and it is the only concept supported by the concept C_j^{LK}, then the probability of knowing the concept C_j^{LK} which is in the set of DKS of A would have the same probability of A, minus a portion of the indirect tested concept error value. Thus, $P(C_j^{LK} | A \in DKS_A) = P(A) - P(e)$. "e" is a portion of the error.

2) If the related (supported) tested concept A is evaluated as a Verified Known Skill and there are other supported concepts by the concept C_j^{LK}, then the probability of knowing the concept such that C_j^{LK} should be calculated by Bayes' Theorem. Using Bayes' Theorem is important for the assumption of the conflicted evaluations of the supported concepts by C_j^{LK}.

3) If the related supported tested concept A is evaluated as a Verified Known Unknown Skill (VNS) and A is the only supported tested concept by the concept C_j^{LK}, then the conditional probability of knowing the concept C_j^{LK}, $P\left(C_j^{LK}|A\right)$, which is in the set of DS of A, would be calculated by Equation 9 and the unconditional probability of concept inside the support set of A which is$P(C_j^{LK})_{DS(A)}$ is changing based on the number of the concepts in the support set of the concept A. This is suggested intuitively.

The Equation 7

$$P(C_j^{LK}{}_{DS(A)}|C_j^{LK}{}_A) = P\left(C_j^{LK}\right) * P(C_j^{LK})_{DS(A)} ,$$

where

$P(C_j^{LK}{}_{DS(A)}|C_j^{LK}{}_A)$ is the conditional probability of knowing the concept C_j^{LK} which is an element in DS of the concept A on a condition of A.

$P(C_j^{LK})_{DS(A)}$ is the unconditional probability of knowing the concept C_j^{LK} and calculated by Equation 10 as:

$$P(C_j^{LK})_{DS(A)} = \frac{1}{\text{Total number of the concepts in DS of A,}\ \ DS(A)}$$

Thus, Equation 8 is

$$P(C)_{DS(A)} = \frac{1}{|DS_{(A)}|}$$

If the concept C_j^{LK} in DNS is a prerequisite only to one tested concept and it is the one concept in DNS of A, then $P(C_j^{LK}|C \in DNS_A) = P(C)_{DS(A)} = P(C_A)$ and will take the same probability of the supported tested concept minus error value of indirectly tested concept.

A. For example, in Figure 5.15, the concept C_7^{LK} supports/prerequisite to only one concept C_3^{LK} and there is another concept C_6^{LK} in the support set of concepts C_3^{LK}. In case of a wrong answer to question q_2, the probability of knowing the concept as a concept supports the tested concept C_3^{LK}, $P(C_7^{LK})$ would be the same probability of C_3^{LK} multiplied by 0.5. Thus, $P(C_7^{LK}|Q_2) = 0.1*0.5= 0.05$,

   where $P(C_7^{LK}) = \frac{1}{|DS_{(A)}|} = \frac{1}{2} = 0.5$ From Equation 10

B. Another example is illustrated in Figure 5.15, where the concept C_4^{LK} which is a support concept to only one concept which is C_1^{LK}, and two other concepts C_5^{LK} and C_6^{LK} are in the set DNS of the concept C_1^{LK}, then the probability of knowing the concept C_4^{LK} will be calculated by Equation 9

   $P(C_4^{LK}{}_{DS(c1)}|\, C_1^{LK}) = P(C_1^{LK})*\, P(C_4^{LK})_{DS(A)}$

   $P(C_4^{LK})_{DS(A)} = \frac{1}{|DS_{(A)}|}$ From Equation 10

   $\mathbf{P}(C_4^{LK})_{\mathbf{DS(A)}} = \frac{1}{3} = 0.33$

Thus,

$P(C_4^{LK}{}_{DS(c1)}|\ C_1^{LK}) = 0.5 * 0.33 = 0.165 = 0.2$

The suggested solution could be proved by using basic formula of Bayes' Theorem to calculate $P(C_4^{LK}|\ C_1^{LK})$, the probability of knowing the concept C_4^{LK} in Example 5.4

By using the example of the concept C_4^{LK}

$$P(C_4^{LK}{}_{DS(c1)}|\ C_1^{LK}) = \frac{P\left(C_1^{LK}\ \middle|C_4^{LK}{}_{DS(c1)}\right)*P(C_4^{LK})}{P(C_1^{LK})} \qquad \text{From Equation 1}$$

$$= \frac{0.5*0.33*0.5}{0.5} = 0.165 = 0.2,$$

where $P\left(C_1^{LK}\ \middle|C_4^{LK}{}_{DS(c1)}\right) = p(C_1^{LK}) * P\left(C_4^{LK}\ \middle|C_1^{LK}{}_{DS(C1)}\right) = 0.5 * 0.5$

$$P(C_4^{LK}) = \frac{1}{\text{Total number of the concepts in DS of } C_1,\ DS(C_1)} \qquad \text{From Equation 10}$$

As shown, the probability of the supported concept C_1^{LK} is presented in both the numerator and the denominator and cancels upon simplification. Thus, the probability of knowing the concept C_4^{LK} is $P(C_4^{LK})=$

$$P(C_1^{LK}) * \frac{1}{\text{Total number of the concepts in DS of } C_1,\ DS(C_1)} \qquad \text{From Equation 10}$$

$P(C_4^{LK}) = 0.5*0.33$

$= 0.165 = 0.2$

Thus, the probability of knowing the concept C_j^{LK} in the prerequisite set of tested concept A, where there are other concepts in the support set of A, would be calculated by the equation Equation 9.

The Equation 9 is

$P(C_j^{LK}{}_{DS(A)}|C_A) = P(C_A) * P(C_j^{LK})_{DS(A)}$

4. If the concept C_j^{LK} in DS is a prerequisite for many tested concepts, then the probability of knowing the concept C_j^{LK} will be calculated using the extended Bayes' Theorem.

A. For example, in Figure 5.15, the concept C_6^{LK} which is a prerequisite concept to the concepts C_1^{LK} C_2^{LK} and C_3^{LK}, will have

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B|A) * P(A) + P(B|\bar{A}) * P(\bar{A})}$$ From Bayes' Theorem

And then

$$P(C_6^{LK}|R) = \frac{P(R|C_6^{LK}) * P(C_6^{LK})}{P(R|C_6^{LK}) * P(C_6^{LK}) + P(R|\bar{C}_6^{LK}) * P(\bar{C}_6^{LK})},$$

Where $R = \{C_1^{LK}, C_2^{LK}, C_3^{LK}\}$. Thus,

$P(R|C_6^{LK}) = P(C_1^{LK}|C_6^{LK}) * P(C_2^{LK}|C_6^{LK}) * P(C_3^{LK}|C_6^{LK})$

$P(C_1^{LK}|C_6^{LK}) = P(C_1^{LK}) * P(C_6^{LK})$ by intuitive which observed from the evaluation of concept 4

$$P(C_6^{LK})_{DS(A)} = \frac{1}{\text{Total number of the concepts in DS of A, DS(A)}}$$ From Equation 10

$$P(C_6^{LK})_{DS(A)} = P(DS)_{(A)} = \frac{1}{|DS_{(A)}|}$$

By replacing concept $C_{\mathbf{6}}^{LK}$ in the Equation 10

$$P(C_6^{LK})_{DS(c_1)} = \frac{\mathbf{1}}{|\mathbf{DS_{c1}}|}$$

$P(C_6^{LK})_{DS(c_1)} = \mathbf{\frac{1}{3}}$

Thus,

$P(C_1^{LK} | C_6^{LK}) = 0.5*0.33 = 0.165$ From Equation 9

Also,

$P(C_2^{LK} | C_6^{LK}) = P(C_2^{LK}) * P(C_6^{LK}) = 0.5 * 0.5 = 0.25$ From Equation 9

$P(C_3^{LK} | C_6^{LK}) = P(C_3^{LK}) * P(C_6^{LK}) = 0.1 * 0.5 = 0.05$ From Equation 9

$P(R | C_6^{LK}) = 0.165 * 0.25 * 0.05 = 0.06*0.13*0.45 = 0.002$

The same equation is applied to calculate the conditional probability of knowing the data R given unknowing the concept C_6^{LK}

$P(R | \bar{C}_6^{LK}) = P(C_1^{LK} | \bar{C}_6^{LK}) * P(C_2 | \bar{C}_6^{LK}) * P(C_3^{LK} | \bar{C}_6^{LK})$

$P(R | C_6^{LK}) = 0.5 * 0.33 * 0.5 * 0.5 * 0.9 * 0.5 = 0.02$

By replacing the result in Equation 2

$P(C_6^{LK} | R) = \frac{0.002}{0.002+0.02} = 0.1$

As part of the same example, the concept $C_{\mathbf{5}}^{LK}$ which is a prerequisite concept to the concepts C_1^{LK}, C_2^{LK}, will have the equation concluded from Bayes' Theorem

$P(\mathbf{C_5^{LK}} | R) = \frac{P(\mathbf{R} | \mathbf{C_5^{LK}}) * P(C_5^{LK})}{P(\mathbf{R} | \mathbf{C_5^{LK}}) * P(C_5^{LK}) + P(\mathbf{R} | \mathbf{\bar{C}_5^{LK}}) * P(\bar{C}_5^{LK})}$ From Equation 2

$R = \{C_1^{LK}, C_2^{LK}\}$,

Thus,

$P(R | C_{\mathbf{5}}^{LK}) = P(C_1^{LK} | C_{\mathbf{5}}^{LK}) * P(C_2^{LK} | C_{\mathbf{5}}^{LK})$

$P(R|C_5^{LK}) = 0.5 * 0.33 * 0.5 * 0.5 = 0.165*0.25 = 0.0415$

also,

$P(R|\bar{C}_5^{LK}) = P(C_1^{LK}|\bar{C}_5^{LK}) * P(C_2^{LK}|\bar{C}_5^{LK})$

$(0.5 * 0.33) * (0.5 * 0.5) = 0.165*0.25 = 0.04$

By replacing the result in above Bayes' Theorem

$$P(C_5^{LK}|R) = \frac{0.04}{0.04+0.04} = \frac{0.04}{0.08} = 0.5$$

Table 5.4 Summary of the Probability Computation Result of Indirectly Tested Concepts in Example 5.4 (Indirectly Tested Concepts are C_4^{LK}, C_5^{LK}, C_6^{LK}, C_7^{LK})

Concept in DS	Related tested concepts	$P(C_j^{LK})$	Number concepts in the DS	$P(\bar{C}_j^{LK})$ Eq3	$P(R\|C_j^{LK})$ Eq6*$P(C_j^{LK})$	$P(R\|\bar{C}_j^{LK})$ Eq7	$P(C_j^{LK}\|R)$
C_4^{LK}	C_1^{LK}	0.5	3	0.5	0.5*0.33=0.165	_	0.2
C_5^{LK}	C_1^{LK}, C_2^{LK}	0.5, 0.5	3,2	0.5	$0.5 * 0.33 + 0.5 * 0.5$ = 0.165+0.25 = 0.415	$0.5 * 0.33 + 0.5 * 0.5$ = 0.165+0.25 = 0.415	0.5
C_6^{LK}	C_1^{LK}, C_2^{LK}, C_3^{LK}	0.5, 0.5, 0.1	3,2,2	0.5	0.165+0.25+0.05 = 0.465	0.165+0.25+0.45 = 0.865	0.1
C_7^{LK}	C_3^{LK}	0.1	1	0.5	0.1*0.5=0.05	_	=0.05

5.6 The Probability of Knowing the Concepts According to the Type of Concept States

The estimation of the right value of the unconditional probability of knowing a concept is a dominant rule in the calculation of the conditional probabilities of knowing concepts in the related Concept Space. There are many cases of the estimation of the unconditional probability of knowing the concept C_j^{LK}, $P(C_j^{LK})$ depends on the Concept State and whether the concept is directly tested, indirectly tested, or has never been tested.

5.6.1 Untested concepts

If the concept C_j^{LK} has never been tested either directly or indirectly such as a concept in DS, then the unconditional probability $P(C_j^{LK})$ will calculated by Equation 5:

$$P\left(C_j^{LK}\right)_{DKS} = \frac{1}{2}$$

Directly Tested Concept (The Concept ∈ VS)

1. Case 1: if the answer to the question asked about the concept C_j^{LK} is correct

and there is only one question asked about the concept, then the tested concept C_j^{LK} is evaluated to be in the set of the concepts Verified Known Skill, and $P(C_j^{LK}) \approx 1$.

2. Case 2: If there are many questions asked about the concept C_j^{LK}, then the unconditional probability of knowing the concept will be Equation 3.

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

Thus,

$$P(C_j^{LK}) = \frac{|Q|}{|Q_q|}$$

The precise probability of knowing the concept C_j^{LK} will be calculated based on Bayes' Theorem

3. Case 3: if the answer to the question asked about the concept C_j^{LK} is incorrect, then the unconditional probability of not knowing the concept will be the number of incorrect responses to the questions testing the concept divided by the total number of questions testing that concept.

The Equation 4

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

Thus,

$$P(\bar{C}_j^{LK}) = \frac{|\bar{Q}|}{|Q_q|},$$

where $\bar{Q}$ is the set of the incorrect answers to the questions asked about the concept C_j^{LK}. The precise probability of knowing the concept C_j^{LK} will be calculated based on Bayes' Theorem (Equation 2). If there is only one question asked about the concept, then $P(\bar{C}_j^{LK}) \approx 1$

5.6.2 Indirectly Tested concept (The Concept ∈ DS)

If the concept C_j^{LK} is an element in DS, it means it is indirectly tested. Particularly, it is a member of the support set of a directly tested concept. Let this directly tested concept be A. Then, the unconditional probability $P(C_j^{LK})$ could have two cases:

1. The concept C_j^{LK} is an element in DKS

The concept C_j^{LK} is a member of derived known set (DKS), which means it was estimated by a question indirectly asked about the concept C_j^{LK}. The probability of knowing the concept C_j^{LK}, $P(C_j^{LK}|A)_{\text{DKS(A)}}$, will be equal to the probability of knowing the supported tested concept minus a portion of error if the evaluator want to consider the error. Thus, $P(C_j^{LK})_{\text{DKS(A)}} \backsimeq 1$.

This equation is proved by the experiment, which proves the validity of the set DKS. This means that if a concept such as C_j^{LK} is a prerequisite to a tested concept, and the tested concept was given, then the prerequisite concept set must be known.

In the case of that the concept $C_j^{LK} \in$ DS which consists of many concepts with contradiction evaluations, which means that it supports many tested concepts, then the probability of knowing the concept such C_j^{LK} will be calculated using Bayes' Theorem. The unconditional probability of knowing the concept C_j^{LK} will be equal between knowing and not knowing the concept = 0.5.

2. The concept C_j^{LK} is an element in DNS; $C_j^{LK} \in$ DNS

If a concept C_j^{LK} at a certain skill level is a member of Known Unknown concepts (DNS) and it has never been directly tested, then the probability of knowing the concept C_j^{LK} depends on three cases:

A. The concept C_j^{LK} supports only one tested concept A, and it is the only concept in the support set of the concept A.

If the concept C_j^{LK} has never been directly tested and it is a concept in the set DNS of one concept A, and the concept C_j^{LK} is the only concept in DNS of the concept A, then the probability of knowing the concept C_j^{LK}, $P(C_j^{LK})$ will take the same value of the concept A minus error value (If the evaluator decides to consider an error value). Thus, $P\left(C_j^{LK} | A \in DNS_A\right) = P(A) - P(e).$ "e" is a portion of the error. The error value, because the concept C_j^{LK} is indirectly tested and evaluated, is based on the supported tested concept A.

B. The concept C_j^{LK} supports only one tested concept, and there are other concepts in the support set of the supported concept A.

If the concept C_j^{LK} has never been directly tested and it is a concept in the set DNS of one concept A, and there are other concepts in DNS of the concept A, then there is unconditional probability of knowing the concept C_j^{LK},

$P(C_j^{LK})_{DS(A)}$ is the unconditional probability of knowing the concept C_j^{LK}. It calculated by Equation 10

$$P\left(C_j^{LK}\right)_{DS(A)} = \frac{1}{\text{Total number of the concepts in the support set of A, DS(A)}}$$

Therefore, the probability of knowing the concept C_j^{LK} on a condition of the supported concept A. $P(C_j^{LK}|A)$ will be calculated using the multiplication of the probability of knowing the supported tested concept A and the probability of knowing the concept C_j^{LK}. The used equation is Equation 9:

$$P(C_j^{LK}{}_{DS(A)}|A) = P(A) * P(C_j^{LK}{}_{DS(A)})$$

For example, let A be a supported concept by C_j^{LK} and there be another concept in the support set of A, and p(A) = 0.1, then the conditional probability of knowing the prerequisite concept C_j^{LK}, would be $P(C_j^{LK})$ = 0.1* 0.5.

Also, the calculation of the concept $C_4^{LK}$and the concept C_7^{LK} in Example 5.4 illustrates the case.

C. The concept C_j^{LK} in DS is a prerequisite for many tested concepts

If the concept C_j^{LK} in DS, is a prerequisite (support) of many tested concepts, then the probability of knowing the concept C_j^{LK} will be calculated by the extended formula of Bayes.

For example, the concept C_6^{LK} which is a prerequisite concept to the concepts C_1^{LK}, C_2^{LK} and C_3^{LK}. Also, for the same case, in the same example, the concept C_5^{LK} which is a prerequisite concept to the concepts C_1^{LK}, C_2^{LK}.

The equation which has been concluded from the extended formula of Bayes. The Equation 2.

$$P\left(C_j^{LK} \middle| R\right) = \frac{P\left(R \middle| C_j^{LK}\right) * P\left(C_j^{LK}\right)}{P\left(R \middle| C_j^{LK}\right) * P\left(C_j^{LK}\right) + P\left(R \middle| \bar{C}_j^{LK}\right) * P\left(\bar{C}_j^{LK}\right)}$$

If there is no previous information about the probability of knowing the concept C_j^{LK}, then the probability of knowing the concept which is in DNS, then the used equation is Equation 9:

$$P\left(C_j^{LK}\right) = P(DNS)_{(Ai)} = \frac{1}{|DNS_{Ai}|}$$

The variable i is to distinguish between the supported concept by the concept C_j^{LK} and indicates the index of the supported concept.

Thus, the equation to calculate the unconditional probability of knowing the concept C_j^{LK} in DS is a prerequisite for many tested concepts is Equation 11.

5.6.3 The Concept ∈ PS

If the concept C_j^{LK} is a member of PS, which means it has never been tested either directly or indirectly, but could be estimated by the related tested concepts either directly or indirectly. If the related concepts, which support concept C_j^{LK} (prerequisite concepts), and all of them are in VS or DS, then the probability $P(C_j^{LK})$ could have two cases:

1. The concept C_j^{LK} is a member of PKS

If the concept C_j^{LK} is a member of the Potential Known Skill set (PKS), this means it is estimated by its support set and all the concepts in its support set at a certain skill level are in either the set VKS or DKS. Let the support set of the concept C_j^{LK} be R. The equation to estimate the probability of knowing the concept C_j^{LK} depends on the probabilities of the support concepts to the concept C_j^{LK}, where the probability of each concept in R □1. The equation which has been concluded from the extended formula of Bayes.

$$P\left(C_j^{LK} \middle| R\right) = \frac{P\left(R \middle| C_j^{LK}\right) * P\left(C_j^{LK}\right)}{P\left(R \middle| C_j^{LK}\right) * P\left(C_j^{LK}\right) + P\left(R \middle| \bar{C}_j^{LK}\right) * P\left(\bar{C}_j^{LK}\right)},$$

where $P(C_j^{LK})$ is the unconditional probability of knowing the concept C_j^{LK}, $P(C_j^{LK}) = 0.5$. Because the concept C_j^{LK} has never been tested, it is given the unconditional probability of half equally knowing and not knowing.

2. The concept C_j^{LK} is a member in PNS

If the concept C_j^{LK} is a member of Potential Known Unknown set (PNS), this means it is estimated by its support set and all the concepts in its support set at a certain skill level are in either the set VNS or DNS. Let the support set of the concept C_j^{LK} be R. The equation to estimate the probability of knowing the concept such C_j^{LK} will depend on R and $P(R) \simeq 0$. The equation which has been concluded from the extended formula of Bayes, which is Equation 2 :

$$P\left(C_j^{LK} \middle| R\right) = \frac{P\left(R \middle| C_j^{LK}\right) * P\left(C_j^{LK}\right)}{P\left(R \middle| C_j^{LK}\right) * P\left(C_j^{LK}\right) + P\left(R \middle| \bar{C}_j^{LK}\right) * P\left(\bar{C}_j^{LK}\right)},$$

where, $P(C_j^{LK})$ is the unconditional probability of knowing the concept C_j^{LK}, $P(C_j^{LK}) = 0.5$. Because the concept C_j^{LK} has never been tested, then it is given that the unconditional probability of half equally knowing and not knowing the concept C_j^{LK}. The equation is Equation 5:

$$P(C_j^{LK})_{PS} = \frac{1}{2}$$

5.7 Assumption of the Values of the Probability of Errors m_r, g_r.

Throughout this dissertation, we assumed a value of probability "e" refers to any kind of errors, such as the lucky guess, the mistake or the type of question, which could be a direct question or an indirect question asked about the concept skill at certain skill level C_j^{Lk}. We assume the values of the probability of the response to question q_r as the following: if the answer Q_r to a question q_r is correct, $P(Q_r|C_r) = 1-g_r$, and $P(\overline{Q}_r|C_r)= m_r$ if the answer $\overline{Q}_r$ to a question $\bar{q}_r$ is wrong. The two constants, $m_r, g_r \in [0,1[$, are respectively called error (careless) probability and guessing probability at q_r. For example, in multiple choice questions, the error probability is high and the evaluator may choose it as one (1) over the number of choices. I did not consider this relation in this study, because the purpose of the errors is to test the equation validation and assign custom errors values based on the number of choices that may affect the result of the study. I assumed specific values for these probabilities to illustrate and test the equations. The result indicated these equations are good to use. I shall now discuss some cases of errors occurring in learner responses to question q_r.

5.7.1 First Suggestion of the Probability of Error:

Let us estimate the probability of occurring errors based on the Concept States and the contradiction in the responses to the set of questions asked about the concept C_j^{LK} at identical skill level.

A. For Verified Skills (VS), which are directly tested concepts.

If the concept is a member in Verified Skills means the concept is directly tested, then the cases will be as the follows:

Case 1: The evaluation of the concept is the result of a response to the question q asked directly about the concept C_j^{LK} at a certain skill level and the question is open-ended, which means it is not multiple choice or true and false questions, then the value of response will be given 1 for a correct answer and 0 for an incorrect answer.

If the question is multiple-choice, then let the probability of error be 0.03. This is just a suggestion, but the error value is depending on the evaluator, it could be 1 divided by the number of choices.

B. For Verified Skills (VS) with Contradiction.

Case 2: There is a contradiction between two questions asked directly about the concept C_j^{LK} at an identical skill level. Then, the probability of the errors is increased. Suppose it is 0.1.

C. For Verified Skills (VS) with contradiction with Derived Skills (DS).

Case 3: There is a contradiction between two questions asked about the concept at the same skill level such that one of them, i.e., q_1, asked directly about the concept and the other q_2 asked about the inference of the evaluation of the concept, such as Derived Skills (DS). Then the probability of the errors m_r, g_r will be assumed 0.1 in the evaluation of the response to question q_1, and it will be higher in q_2, which asked about the concept by inference. Let the values of the probability of errors associated with the evaluation of the responses to the questions, such as q_2, be 0.2. Table 5.5 illustrates the proposed assumption values of the probability of errors m_r and g_r

Table 5.5 Summary of the Proposed Assumption Values of the m_r and g_r Based in First Suggestion of Error Value

Type of q	Response	Type of Concept State	Contradiction	Probability of careless error m_r	Probability of lucky guess g_r	The response probability
Multiple Choice q	Correct	VKS/DKS	No	-	0.03	$P(Q_r \mid C) = (1-0.03) = 0.97$
Multiple Choice q	Incorrect	VNS/DNS	No	0.03	-	$P(\overline{Q}_r \mid C) = 0.03$
Multiple Choice q	Correct	VKS	Yes =0.1	-	0.03+0.1=0.13	$P(Q_r \mid C) = (1-0.13) = 0.87$
Multiple Choice q	Correct	DKS	Yes = 0.1	-	0.03+0.2=0.23	$P(Q_r \mid C) = (1-0.23) = 0.77$
Multiple Choice q	Incorrect	VNS	Yes = 0.1	0.03+0.1=0.13	-	$P(\overline{Q}_r \mid C)=0.13$
Multiple Choice q	Incorrect	DNS	Yes = 0.2	0.03+0.2=0.23	-	$P(\overline{Q}_r \mid C) = 0.23$
Regular q	Correct	VKS/DKS	No	0	0	$P(Q_r \mid C) =1$
Regular q	Incorrect	VNS/DNS	No	0	0	$P(\overline{Q}_r \mid C)=0$
Regular q	Correct	VKS	Yes =0.1	-	0.1	$P(Q_r)$ =(1-0.1)

						= 0.9
Regular q	Correct	DKS	Yes= 0.1	-	0.2	$P(Q_r\|C)$ = $(1-0.2) = 0.8$
Regular q	Incorrect	VNS	Yes= 0.1	0.1	-	$P(\overline{Q}_r\|C)$=0.1
Regular q	Incorrect	DNS	Yes = 0.1	0.2	-	$P(\overline{Q}_r\|C)$ =0.2

5.7.2 Example 5.5: The Assumption of Error Values Based on the Type of Question and the Concept State

1. **Case 1**

Let a direct question q_1 be a multiple-choice question asked about the concept C_1^{LK} at skill level 3 denoted as C_1^{L3} .

If the analysis result of the grader gives an evaluation such that the response r of a learner S_1 is a correct answer to the question q_1, then g_1 = 0.03 Thus, $p(Q_1|C)$ = $1-0.03 = 0.97$.

2. **Case 2**

Let a direct question q_2, be not a multiple-choice question asked about the concept $C_1^{(L3)}$.

If the analysis result of the grader gives an evaluation such that the response Q_1 of a learner S_1 is a correct answer to the question q_1, then m_1, g_1 = 0. Thus, $p(Q_1|C)$ = 1.

If the grader gives an evaluation such that the response $\overline{Q}_2$ of a learner S_1 is incorrect answer to question q_2, which means there is a contradiction with Q_1, then $m_2 =$ 0.1 and g_1 = 0.1 + 0.03= 0.13 or if the evaluator prefers to not consider the contradiction and the probability of knowing, then the concept will be calculated by the Bayes' Theorem, since it counts the product of all the evaluations and the number of correct responses over all the questions asked about the concept.

3. Case 3

Let there also be a question q_3 asked, which infers the knowing of the concept such that the concept C_1^{L3} is a member of DS.

If the analysis result of the grader gives an evaluation such that the response Q_3 of a learner S_1 is a correct answer to the question q_3, which means there is a contradiction with q_2 then $g_3 = 0.2$ and $m_2 = 0.1$.

Thus, $P(Q_3|C) = (1-g_3) = (1-0.2) = 0.8$, and $P(\overline{Q}_2|C) = 0.1$.

If the analysis result of the grader gives an evaluation such that the response $\overline{Q}_3$ of a learner S_a is an incorrect answer to the question q_3, it means there is a contradiction with q_1 then $m_3 = 0.2$ and $g_1 = 0.1 + 0.03 = 0.13$. The value 0.03 is a portion of the multiple-choice type of the question q_1. Thus, $P(\overline{Q}_3|C) = (1 - m_3) = (1-0.2) = 0.8$.

Let's summarize the case of a contradiction in the set of evaluations of responses to the set of questions asked about a concept at identical skill level.

Let the data given about a learner S_1 be the following:

Set of questions Q_q asked about concept C_1 at level 3 which denoted as C_1^{L3} is $\{q_1,q_2,q_3\}$. "q_1" is multiple choice, q_2 is open-ended regular question, such as an essay question directly asked about concept C_1^{L3} and q_3 is a regular (Open Ended) question asked about C_1^{L3} by inference. Thus, by question q_1 and question q_2, the concept C_1^{L3} is a member of

VS, and by q_3 is a member of DS. The set of the probabilities of evaluations of the responses $R = \{Q_1, \overline{Q}_2, Q_3\}$,

where Q_1 is a correct answer, $\overline{Q}_2$ is an incorrect answer and Q_3 is a correct answer.

Therefore, the error probabilities values will be as e_{q1} = 0.87, e_{q2} = 0.1 and e_{q3} = 0.8, since q_3 is a question asked about concept C_1^{L3} by inference, thus, R = {0.87, 0.1, 0.8}. Figure 5.16, and Table 5.6 illustrates cases of contradiction of some responses to the questions about the concept C_1^{L3} with different cases of assumption of error values. Also, Table 5.7 shows the computation of the probability of knowing the concept C_1^{L3} based on the contradiction in the responses as in Table 5.6.

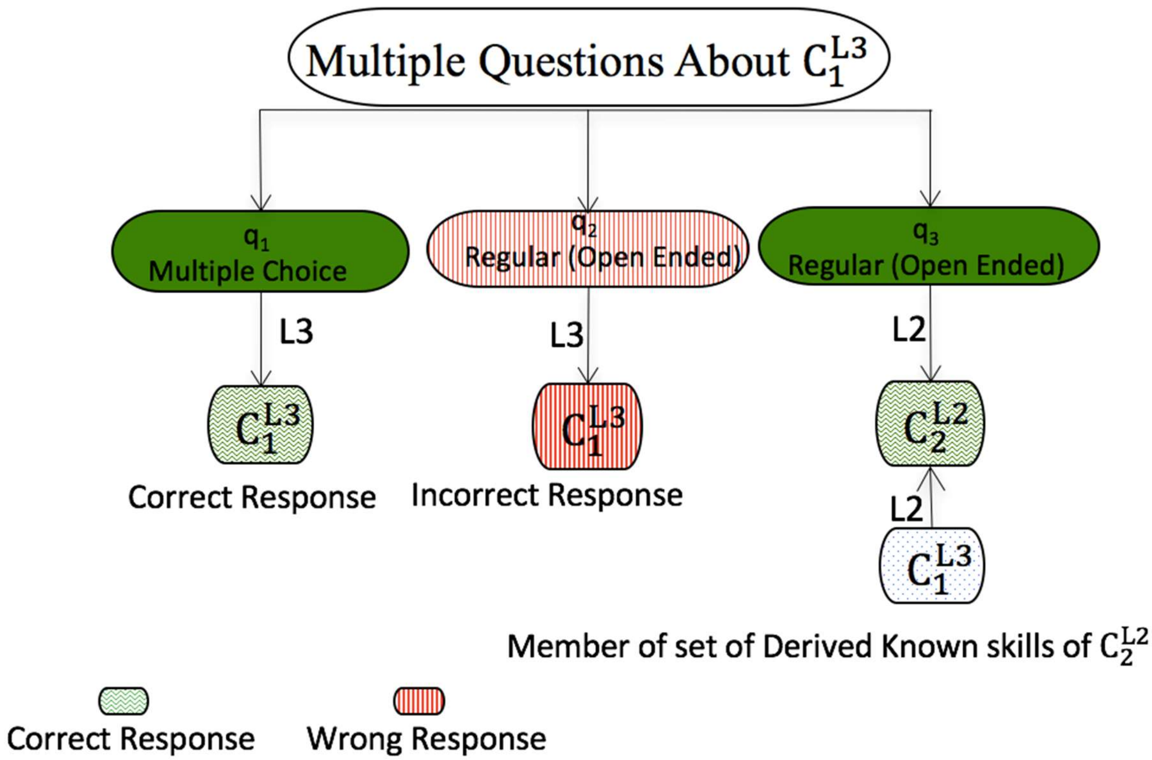


Figure 5.16 An Example of Various Types of Questions about the Concept C_1^{L3}

Table 5.6 The Cases of Contradiction of Some Responses About Concept C_j^{Lk} with Different Cases of Concept State

Q	Type of question	Response	Type of concept state	C_r	g_r	P(R)
q_1	Multiple Choice	1	VKS	-	0.13 Contradiction probability + multiple choice = 0.1+0.03	$p(Q_1\|C) = 0.87$
q_2	Regular question	0	VNS	0.1	-	$p(\bar{Q}_2\|C) = 0.1$
q_3	Regular question	1	DKS	-	0.2 DKS member	$P(Q_3\|C) = 1\text{-}0.2 = 0.8$

Table 5.7 The Calculation of the Probability of Knowing the Concept C_j^{Lk} based on First Suggestion

	$P(C_j^{Lk})$ Eq. 3	$P(\bar{C}_j^{Lk})$ Eq. 4	$P(R\|C_j^{Lk})$ Eq. 6	$P(R\|\bar{C}_j^{Lk})$ Eq. 7	$P(C_j^{Lk}\|R)$ Eq. 2	$P(\bar{C}_j^{Lk}\|R)$ Eq. 8
By using first suggestion (Equation 4 and Equation 5) with Equation 1	2/3=0.67	1/3=0.33	0.07	0.02	0.86	0.14

5.7.3 Second Suggestion of the Probability of Errors

Let us estimate the probability of errors based on the Concept States, and ignore any contradictions in the responses to the set of questions asked about the concept C_j^{Lk} at an identical skill level. We don't consider contradictions, if any, because we consider the values of the errors for all the responses to questions about the concept, even if there is a contradiction in the responses. Any evaluations which contradict each other will reverse each other in the calculation using Bayes' Theorem. In this case, we could give the probability of errors for all the evaluations even if the question is open-ended, i.e., not multiple choice. Let the value of the wrong answer be 0.1 and let the probability value of errors if the Concept State be DS = 0.1. Thus, the probability of error when the question indirect asked about the concept = 0.2, is the probability of error for open-ended question = 0.1 plus the probability of indirect question = 0.1. Table 5.8 shows the responses probabilities without considering the contradiction. The calculation of the probability of knowing the concept C_1^{L3} without considering the contradiction based on as in Table 5.9.

Table 5.8 Summary of the Proposed Assumption Values of the Errors m_r and g_r Based on Second Suggestion

Type of q_r	Response	Type of CS	Contradiction	P($\mathcal{C}_r$)	P(g_r)	Response Evaluation
Multiple Choice q_r	Correct	VKS/DKS	No	-	0.03+0.1	$P(Q_r\|C)$ = (1-0.13)= 0.87
Multiple Choice q_r	Incorrect	VNS/DNS	No	0.03+0.1	-	$P(\bar{Q}_r\|C)$=0.13
Multiple Choice q_r	Correct	VKS	Yes = 0	-	0.03+0.1	$P(Q_r\|C)$ = (1-0.13) = 0.87
Multiple Choice q_r	Correct	DKS	Yes = 0	-	0.03+0.2=0.23	$P(Q_r\|C)$ = (1-0.23) = 0.77
Multiple Choice q_r	Incorrect	VNS	Yes =0	0.03+0.1=0.13	-	$P(\bar{Q}_r\|C)$=0.13
Multiple Choice q_r	Incorrect	DNS	Yes = 0	0.03+0.2=0.23	-	$P(\bar{Q}_r\|C)$=0.23
open-ended q_r	Correct	VKS	NO	0	0.1	$P(Q_r\|C)$ =1-0.1=0.9
open-ended q_r	Incorrect	VNS	No	0.1		$P(\bar{Q}_r\|C)$=0.1
open-ended q_r	Correct	VKS	Yes	-	0.1	$(Q_r\|C)$ =(1-0.1) = 0.9
open-ended q_r	Correct	DKS	Yes/No	-	0.2	$(Q_r\|C)$ = (1-0.2) = 0.8
open-ended q_r	Incorrect	VNS	Yes/No	0.1	-	$P(\bar{Q}_r\|C)$=0.1
open-ended q	Incorrect	DNS	Yes/No	0.2	-	$P(\bar{Q}_r\|C)$=0.2

Table 5.9 Second Suggestion of the Calculation of the Probability of Knowing the Concept C_1^{L3} Based on the Responses as in Table 5.8

	$P(C_j^{Lk})$ Eq. 3	$P(\overline{C}_j^{Lk})$ Eq. 4	$P\left(R \middle\| C_j^{Lk}\right)$ Eq. 6	$P\left(R \middle\| \overline{C}_j^{Lk}\right)$ Eq. 7	$P(C_j^{Lk} \| R)$ Eq. 2	$P(\overline{C}_j^{Lk} \| R)$ Eq. 8
By using second suggestion (equation Eq.6 and Eq.7) with eq1	2/3=0.67	1/3=0.33	(0.87+0.8)*0.67= 1.12	0.1*0.33= 0.03	0.99	0.01

CHAPTER 6

Estimation of the Concept States' Zones According to the Human Subject Test

The test is introduced online in one session. The participants are 154 learners from graduate level, attending the CS 61002 Algorithms and Programming class in the Computer Science Department. The two types of questions (DQ and OQ) were administered to the learners. The OQ are typical of the kind of questions that are usually given to the learners for their midterm exam. We prepared the DQ type of questions from the midterm questions and gave an online test in one session to the participants at the end of the semester. The online test contained 47 questions. The first 9 questions were selected from the OQ, and the remaining 38 questions were direct questions asking about the exact skills that were extracted from the analysis of the OQ. Therefore, each concept is evaluated using at least two questions at the same skill level of the concept. The responses to the questions will form the dataset R. The dataset R includes the set of the response probability to the asked questions about the concept at a certain skill level. The probability of the responses is set such that: if the answer is correct, then the probability of knowing the concept C_j is $P(C_j|Q_r) = (1-g_r)$ and the probability of not knowing the concept C_j is $P(\bar{C}_j|Q) = g_r$. On the other hand, if the response to a question q_r is incorrect, then the probability of knowing the concept C_j, which has been asked by the question q_r is $P(C_j|\bar{Q}) = \mathcal{C}_r$ and the probability of not knowing the concept C_j is $P(\bar{C}_j|\bar{Q}) = (1-\mathcal{C}_r)$. The two constants, $\mathcal{C}_r, g \in [0,1[$, are respectively termed (careless) error probability and guessing probability at q_r.

In the implementation of Bayes' Theorem on the real test, we assume the error values based on many cases.

The total number of the tested concepts without considering the skill levels is 44. In the real exam, some concepts are tested on many levels, while other concepts are tested on only one level. The number of tested concepts with frequencies are 80.

In the computation of the probability of knowing the concepts there are many cases:

Case 1: The direct tested concept will be given the probability of knowing and un-knowing the concept, based on the type of question and the response to the question, as the following:

1) If the response is correct, then the probability of knowing the concept C_j is $P(C_j|Q_r) = (1-g_r) = 1-0.1=0.9$ and the probability of not knowing the concept $\bar{C}_j$ is $P(\bar{C}_j|Q_r) = g_r = 0.1$.
2) If the response to a question q_r is incorrect, then the probability of knowing the concept C_j, which has been asked by the question q_r is $P(C_j|\bar{Q}_r)) = c_r = 0.1$ and the probability of not knowing the concept C_j is $P(\bar{C}_j|\bar{Q}_r)) = (1-c_r) = 0.9$.
3) If the question is multiple-choice and it is a direct question, then let the probability of error be 0.03+0.1 = 0.13.
4) If the concept at a certain skill level is tested by an open question, which is a regular question that has been prepared by the instructor and hasn't been concentrated in the skill level of the concept, then let the error value be 0.2.

5) If the question is multiple-choice and is an indirect question, then let the probability of error be 0.03+0.2 = 0.23.

Case 2: The concept is untested, but evidence is given that it is known or unknown such that the untested concept is a member of the set DS or PS of a tested related concept. The concept which is a member in DS or PS will be given the probability of knowing and un-knowing the concept based on the response to the question that asks about the related (supported) tested concept as the following:

1) If the response to the question asked about the related (supported) concept is correct, then the error value of the estimated probability of knowing the concept will be given a higher value than the error value in the question asked about the supported concept. For example, let "A" be a target (supported) concept, and concept C_j be a member in the support set of A denoted as SS(A). If an error value of the question q_r which has been asked about a concept A is 0.1, then the error value in the estimation of the probability of knowing the concept C_j will be given a higher value such as 0.2. Therefore, the probability of knowing the concept C_j will be $P(C_j) = (1-g_r) = 1 - 0.2 = 0.8$ and the probability of not knowing the concept C_j will be $P(\bar{C}_j) = g_r = 0.2$

2) If the response to the question q_r asked about the target (supported) concept A is incorrect, then the error value of the estimated probability of not knowing the concept C_j is increased. Therefore, the probability of knowing the concept $C_j =$ will

be $P(C_j) = \mathcal{C}_r = 0.2$ and the probability of not knowing the concept C_j will be $P(\bar{C}_j) = (1 - \mathcal{C}_r) = 0.8$

To reveal the precise probability of knowing the untested concept we have to consider the set of all the concepts in the support set of the tested concept. Therefore, Bayes' Theorem will be used as the Equation 2.

$$P(\mathbf{C_j}|R) = \frac{P(\mathbf{R}|\mathbf{C_j}) * P(C_j))}{P(\mathbf{R}|\mathbf{C_j}) * P(C_j) + P(\mathbf{R}|\bar{\mathbf{C}}_\mathbf{j}) * P(\bar{C}_j)} \qquad \text{Equation 2}$$

- C_j denotes knowing the concept C_j

- $P(C_j)$ is the unconditional probability of knowing the concept C_j, which is the initial probability of knowing the concept C_j. It is just the rate of the correct responses to the questions asked about the concept C_j.

Case 3: The concept tested many times, through either direct or indirect evaluation, will be given the probability of knowing and un-knowing the concept based on Bayes' Theorem.

In the implementation of Bayes' Theorem in the real test, we used the Theorem to calculate the probability of each tested concept, since we had at least two responses to the questions asked about the same concept. Thus, the unconditional probability of each concept is assigned the initial value as illustrated in Case 1. Then, Bayes' Theorem may be used to calculate the precise probability based on the set of response data.

The learning object, which is the inclusion of all the concepts in the domain, will be found by calculating the probabilities of VS and VNS; DS and DNS; PS and PNS. The probability of VS and VNS indicate the higher more-complex concepts among the learning object concepts. The probability of DS and DNS indicate the less complicated concepts in the learning objects domain. PKS and PNS indicate the tangential concepts in the domain. Thus, in the implementation of the computation of the validation test using Bayes' Theorem, we classified the tested concepts into three basic proposed sets: VS, DS and PS.

6.1 The Computation of the Probability of Knowing a Concept in VS

Each concept is directly tested by two questions: the question prepared by the instructor and the question prepared based on the analysis of the instructor's questions. The learner response to the question based on a certain skill level of a concept, which is an element in VS, is evaluated as the following:

1) If the response is correct, then the probability of knowing the concept C_j is $P(C_j|Q_r) = (1-g_r) = 1-0.1=0.9$, and the probability of not knowing the concept q_r is $P(\overline{C}_j|Q_r) = g_r = 0.1$. "$g_r$" is a lucky guess error with suggested value = 0.1. r is an integer number that refers to the index of the related question.
2) If the response to a question q_r is incorrect, then the probability of knowing the concept C_j, which has been asked by the question q_r is $P(C_j|Q_r) = \mathcal{C}_r = 0.1$, and the probability of not knowing the concept C_j is $P(\overline{C}_j|\overline{Q}_r) = (1-\mathcal{C}_r) = 0.9$. "$\mathcal{C}_r$" is a careless error with suggested value = 0.1.

In conclusion, we observed that if there are two correct answers to the questions asked about a concept at the same skill level, the result of the probability computation is equal to one, even though we consider the probability of error in each response. The probability of knowing concept C_j, that is evaluated based upon two questions which have conflicted responses, shows that the probability of knowing the concept is changing based on the type of questions. In order to determine the truth estimation, we suggest providing an error value to the responses based on the type of the question. If the question type is multiple choice, then the probability of the error is increased. Also, a higher probability of error is given to the question indirectly asking about the target skill, rather than the question directly asking about the skill. This is proved by the result of the computation of knowing the concepts in a real test. Table 6.1 illustrates the computation of the probability of knowing the concept, which is an element in DS, by using the suggested Bayes' Formula and errors values. Table 6.2 illustrates the computation of the probability of not knowing the concept, which is an element in DS, by using the suggested Bayes' Formula and errors values. Figures 6.1 and 6.2 show the probability of knowing the concepts in the domain of VS according to the evaluation result of the perfect learner and the randomly selected learner (who didn't answer all the questions correctly) respectively. Also, Figure 6.3 shows the probability of 5 learners including the perfect learner.

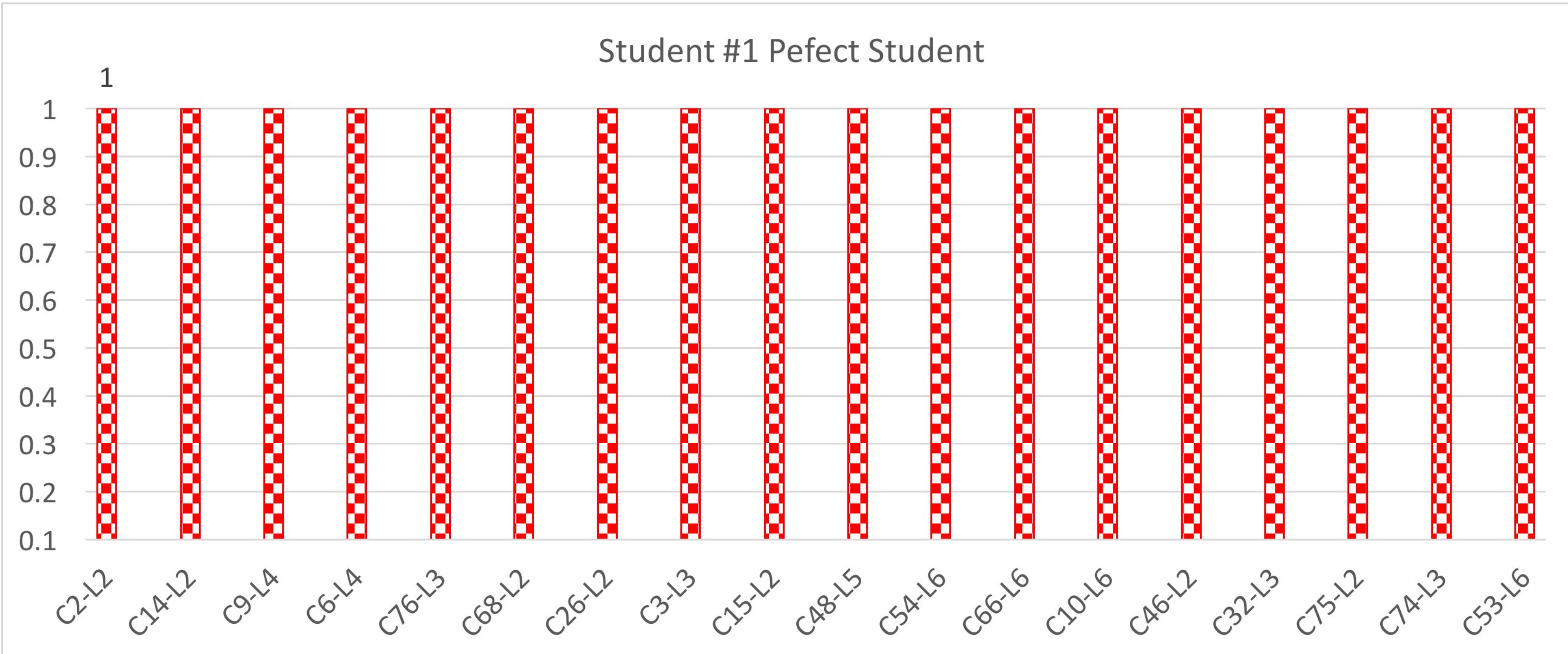
Table 6.1 Summary of Calculation of the Probability of Knowing a Concept in the Set of VS According to Result of Human Subgect Test

	Response to q_1 Q_1	Response to q_2 Q_2	$P(Q_1)$ Knowing	$P(Q_2)$ Knowing	Eq3 $P(C_j)$	Eq6 $P(R\|C_j)$	Eq 2 $P(C_j\|R)$
C 1	1	1	0.9	0.9	1	0.81	1
C 2	0	0	0.1	0.1	0	0.01	0
C 3	1	0	0.9	0.1	0.5	0.09	0.5
C 4	0	1	0.1	0.9	0.5	0.09	0.5

Table 6.2 Summary of Calculation the Probability of Not Knowing a Concept in the Set of VS According to the Result of Human Subject

	Response Q_1	Response Q_2	$P(Q_1)$ Not Knowing	$P(Q_2)$ Not Knowing	Eq4 $P(\bar{C}_j)$	Eq7 $P(R\|\bar{C}_j)$	Eq 8 $P(\bar{C}_j\|R)$
C1	1	1	0.1	0.1	0	0.01	0
C2	0	0	0.9	0.9	1	0.81	1
C3	1	0	0.1	0.9	0.5	0.09	0.5
C4	0	1	0.9	0.1	0.5	0.09	0.5

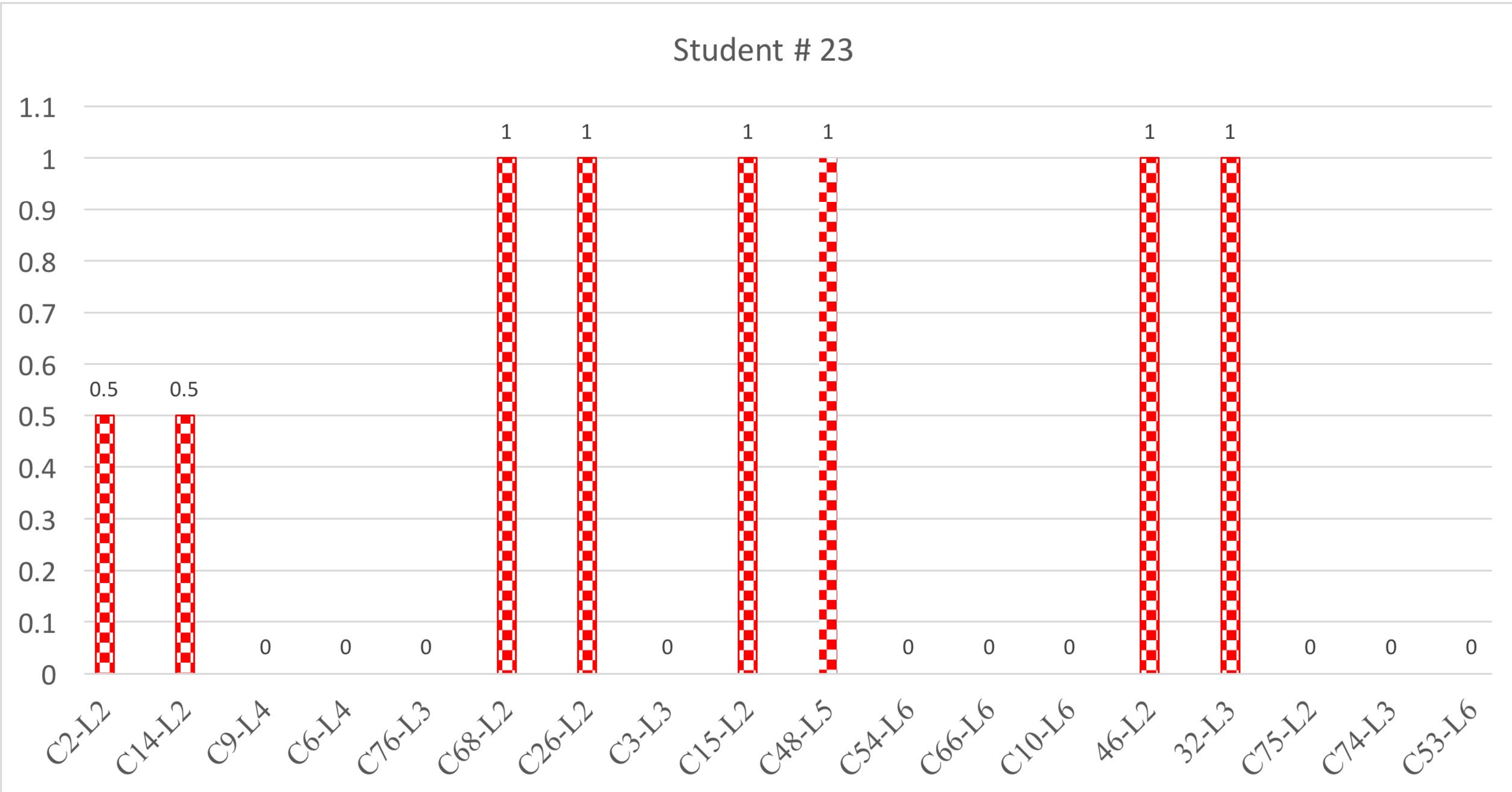
The tables show the changing in the probability of knowing the concept according to the responses to two questions asked about the same skill level of the concept. We assume the probability of occurring the errors in the responses is equal to 0.1 in both the correct response and the incorrect response. If we change the values of errors based on the type of the question, the conditional probability of knowing a concept according to the conflicted responses to the two questions asking about the same skill of the concept will be changed from 0.5. This is used in the computation of the probability of knowing the concept in DS and the PS domain



X : Concept # at certain skill level

Y : The probability of knowing the concept

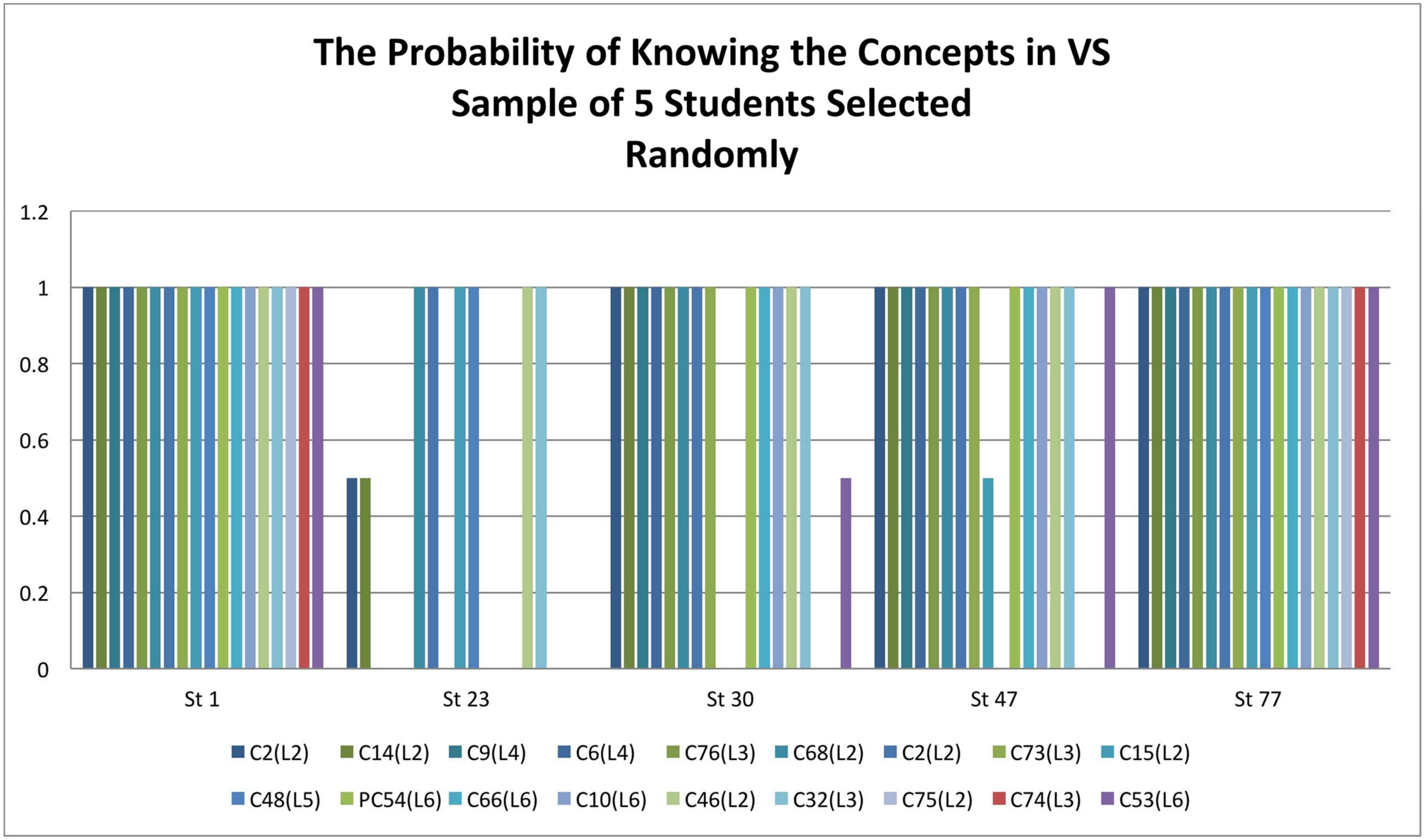
Figure 6.1 The Probability of Knowing the Concepts in VS According to the Evaluation of the Perfect Learner



X : Concept # at certain skill level

Y : The probability of knowing the concept

Figure 6.2 The Probability of Knowing the concepts in VS According to the Evaluation of Randomly Selected Learner



X : Learner Number
Y : Probability of knowing the concept

Figure 6.3 The Probability of Knowing the Concepts in VS

- Sample of 5 Learners Randomly Selected from the 154 Participants

6.2 The Computation of the Probability of Knowing a Concept in DS and PS

In the computation of the probability of knowing or not knowing the concept, both of which are members in either DS or PS, we change the error values based on the type of question. We attribute an unequal value to the direct and indirect question. Also, we increase the error value in the question which is multiple choice.

1) If the concept at a certain skill level is tested by an indirect question, which is an open-ended regular question that has been prepared by the instructor and hasn't been concentrated in the skill level of the concept or asked about the supported concept rather than the evaluated concept, then let the errors value be 0.2

2) If the questions are multiple-choice and open question, then let the error value be 0.03+0.2 = 0.23.

3) If the response is correct, then the probability of knowing the concept C_j is $P(C_j|Q_r)$ = $(1-g_r)$= $1-0.23=0.77$ and the probability of not knowing the concept C_j is $P(\bar{C}_j|Q_r) = g_r = 0.23$. "$g_r$" is a lucky guess error with the suggested value = 0.23.

4) If the response to a question q_r is incorrect, then the probability of knowing the concept C_j, which has been asked by the question q_r is $P(C_j|\bar{Q}_r) = \mathcal{C}_r = 0.23$ and the probability of not knowing the concept C_j is $P(\bar{C}|\bar{Q}) = (1 - \mathcal{C}_r) = 0.77$. "$\mathcal{C}_r$" is careless error with suggested value = 0.23.

Table 6.3 Summary of the Cases of the Estimated Result of the Probability of Knowing a Concept in the Set of DS and PS According to the Real Test

Types of the questions	C #	Response Q_1	Response Q_2	$P(Q_1\|C_j)$	$P(Q_2\|C_j)$	$P(\overline{Q_1}\|\overline{C}_j)$	$P(\overline{Q_2}\|\overline{C}_j)$	Eq. 2 $P(C_j\|R)$	Eq. 8 $P(\overline{C}_j\|R)$
q_1 is indirect question-Essay type q_2 is direct question-Essay type	C 1	1	1	0.8	0.9	0.1	0.1	1	0

Types of the questions	C #	Response Q_1	Response Q_2	$P(Q_1\|C_j)$	$P(Q_2\|C_j)$	$P(\overline{Q_1}\|\overline{C}_j)$	$P(\overline{Q_2}\|\overline{C}_j)$	Eq. 2 $P(C_j\|R)$	Eq. 8 $P(\overline{C}_j\|R)$
q_1 is indirect question-Essay type q_2 is direct question-Essay type	C 2	0	0	0.2	0.1	0.8	0.9	0	1

Types of the questions	C #	Response Q_1	Response Q_2	$P(Q_1\|C_j)$	$P(Q_2\|C_j)$	$P(\overline{Q_1}\|\overline{C}_j)$	$P(\overline{Q_2}\|\overline{C}_j)$	Eq. 2 $P(C_j\|R)$	Eq. 8 $P(\overline{C}_j\|R)$
q_1 is indirect question-Essay type	C 3	1	0	0.8	0.1	0.1	0.8	0.31	0.69

q_2 is direct question-Essay type									

Types of the questions	**C #**	**Response Q_1**	**Response Q_2**	**$P(Q_1\|C_j)$**	**$P(Q_2\|C_j)$**	**$P(\overline{Q_1}\|\bar{C}_j)$**	**$P(\overline{Q_2}\|\bar{C}_j)$**	**Eq. 2 $P(C_j\|R)$**	**Eq. 8 $P(\bar{C}_j\|R)$**
q_1 is indirect question-Essay type q_2 is direct question-Essay type	C4	0	1	0.2	0.9	0.8	0.1	0.69	0.31

Types of the questions	**C #**	**Response Q_1**	**Response Q_2**	**$P(Q_1\|C_j)$**	**$P(Q_2\|C_j)$**	**$P(\overline{Q_1}\|\bar{C}_j)$**	**$P(\overline{Q_2}\|\bar{C}_j)$**	**Eq. 2 $P(C_j\|R)$**	**Eq. 8 $P(\bar{C}_j\|R)$**
q1 is indirect question and Multiple choice q2 is direct question Essay	C5	1	1	0.77	0.9	0.23	0.1	1	0

Types of the questions	C #	Response Q_1	Response Q_2	$P(Q_1\|C_j)$	$P(Q_2\|C_j)$	$P(\overline{Q_1}\|\bar{C}_j)$	$P(\overline{Q_2}\|\bar{C}_j)$	Eq. 2 $P(C_j\|R)$	Eq. 8 $P(\bar{C}_j\|R)$
q1 is indirect question and Multiple choice q2 is direct question Essay	C 6	0	0	0.23	0.1	0.77	0.9	0	1

Types of the questions	C #	Response Q_1	Response Q_2	$P(Q_1\|C_j)$	$P(Q_2\|C_j)$	$P(\overline{Q_1}\|\bar{C}_j)$	$P(\overline{Q_2}\|\bar{C}_j)$	Eq. 2 $P(C_j\|R)$	Eq. 8 $P(\bar{C}_j\|R)$
q_1 is indirect question and Multiple choice q_2 is direct question- Essay	C 7	0	1	0.23	0.9	0.77	0.1	0.73	0.27

Types of the questions	C #	Response Q_1	Response Q_2	$P(Q_1 \mid C_j)$	$P(Q_2 \mid C_j)$	$P(\overline{Q_1} \mid \bar{C}_j)$	$P(\overline{Q_2} \mid \bar{C}_j)$	Eq. 2 $P(C_j \mid R)$	Eq. 8 $P(\bar{C}_j \mid R)$
q_1 is indirect question and Multiple choice q_2 is direct question Essay	C8	1	0	0.77	0.1	0.23	0.9	0.27	0.73

Types of the questions	C #	Response Q_1	Response Q_2	$P(Q_1 \mid C_j)$	$P(Q_2 \mid C_j)$	$P(\overline{Q_1} \mid \bar{C}_j)$	$P(\overline{Q_2} \mid \bar{C}_j)$	Eq. 2 $P(C_j \mid R)$	Eq. 8 $P(\bar{C}_j \mid R)$
q_1 is indirect question- Essay q_2 is direct question- Multiple Choice	C9	1	1	0.8	0.87	0.2	0.13	1	0

Types of the questions	C#	Response Q_1	Response Q_2	$P(Q_1\|C_j)$	$P(Q_2\|C_j)$	$P(\overline{Q_1}\|\bar{C}_j)$	$P(\overline{Q_2}\|\bar{C}_j)$	Eq. 2 $P(C_j\|R)$	Eq. 8 $P(\bar{C}_j\|R)$
q_1 is indirect question-Essay q_2 is direct question-Multiple Choice	C10	0	0	0.2	0.13	0.8	0.87	0	1

Types of the questions	C#	Response Q_1	Response Q_2	$P(Q_1\|C_j)$	$P(Q_2\|C_j)$	$P(\overline{Q_1}\|\bar{C}_j)$	$P(\overline{Q_2}\|\bar{C}_j)$	Eq. 2 $P(C_j\|R)$	Eq. 8 $P(\bar{C}_j\|R)$
q_1 is indirect question-Essay q_2 is direct question-Multiple Choice	C11	1	0	0.8	0.13	0.2	0.87	0.37	0.63

Types of the questions	C#	Response Q_1	Response Q_2	$P(Q_1\|C_j)$	$P(Q_2\|C_j)$	$P(\overline{Q_1}\|\bar{C}_j)$	$P(\overline{Q_2}\|\bar{C}_j)$	Eq. 2 $P(C_j\|R)$	Eq. 8 $P(\bar{C}_j\|R)$
q_1 is indirect question- Essay q_2 is direct question- Multiple Choice	C12	0	1	0.2	0.87	0.8	0.13	0.63	0.37

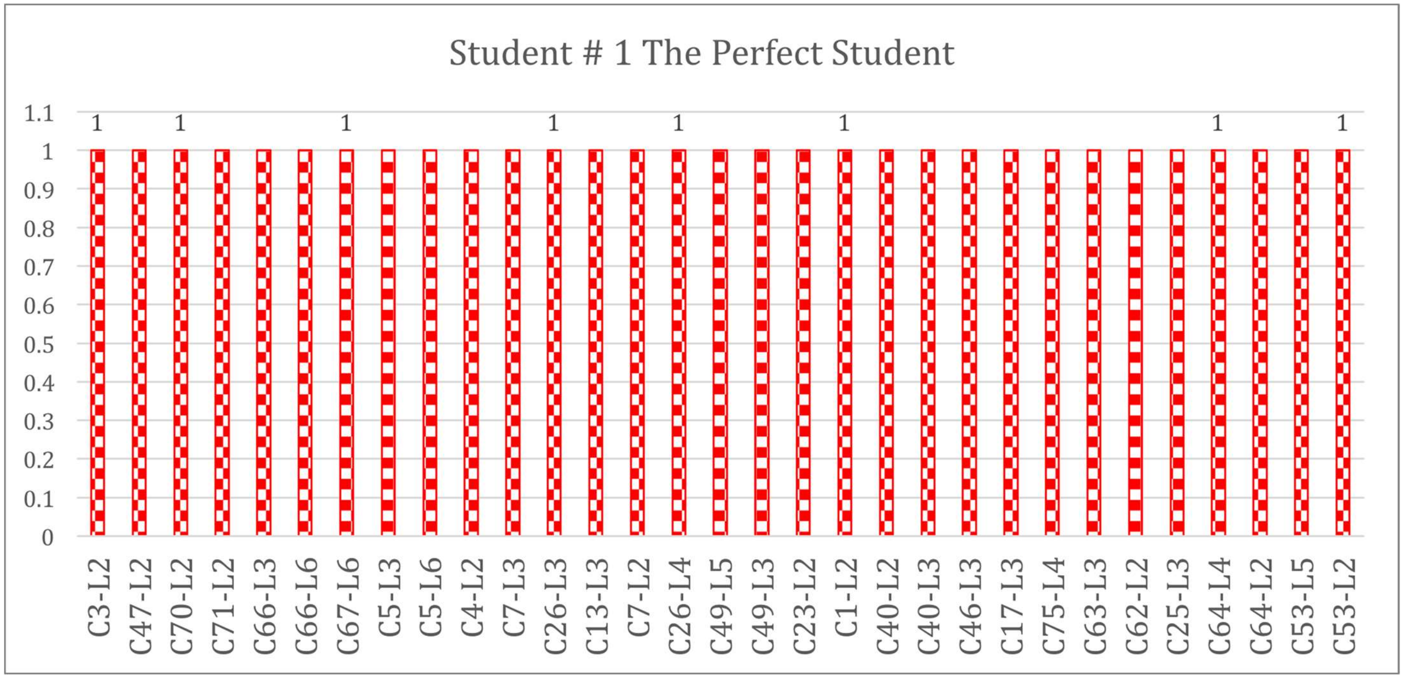
Types of the questions	C#	Response Q_1	Response Q_2	$P(Q_1\|C_j)$	$P(Q_2\|C_j)$	$P(\overline{Q_1}\|\bar{C}_j)$	$P(\overline{Q_2}\|\bar{C}_j)$	Eq. 2 $P(C_j\|R)$	Eq. 8 $P(\bar{C}_j\|R)$
q_1 is indirect question and multiple choice q_2 is direct question and multiple choice	C13	1	1	0.77	0.87	0.23	0.13	1	0

Types of the questions	C#	Response Q_1	Response Q_2	$P(Q_1 \vert C_j)$	$P(Q_2 \vert C_j)$	$P(\overline{Q_1} \vert \bar{C}_j)$	$P(\overline{Q_2} \vert \bar{C}_j)$	Eq. 2 $P(C_j \vert R)$	Eq. 8 $P(\bar{C}_j \vert R)$
q_1 is indirect question-Multiple Choice q_2 is direct question-Multiple Choice	C14	0	0	0.23	0.13	0.77	0.87	0	1

Types of the questions	C#	Response Q_1	Response Q_2	$P(Q_1 \vert C_j)$	$P(Q_2 \vert C_j)$	$P(\overline{Q_1} \vert \bar{C}_j)$	$P(\overline{Q_2} \vert \bar{C}_j)$	Eq. 2 $P(C_j \vert R)$	Eq. 8 $P(\bar{C}_j \vert R)$
q_1 is indirect question-multiple choice q_2 is direct question-multiple choice	C15	1	0	0.77	0.13	0.23	0.87	0.33	0.67

Types of the questions	C#	Types of the questions	C#	Response Q_1	Response Q_2	$P(Q_1 \mid C_j)$	$P(Q_2 \mid C_j)$	$P(\overline{Q_1} \mid \bar{C}_j)$	$P(\overline{Q_2} \mid \bar{C}_j)$
q_1 is indirect question- multiple choice	C16	0	1	0.23	0.87	0.77	0.13	0.67	0.33
q_2 is direct question- multiple choice									

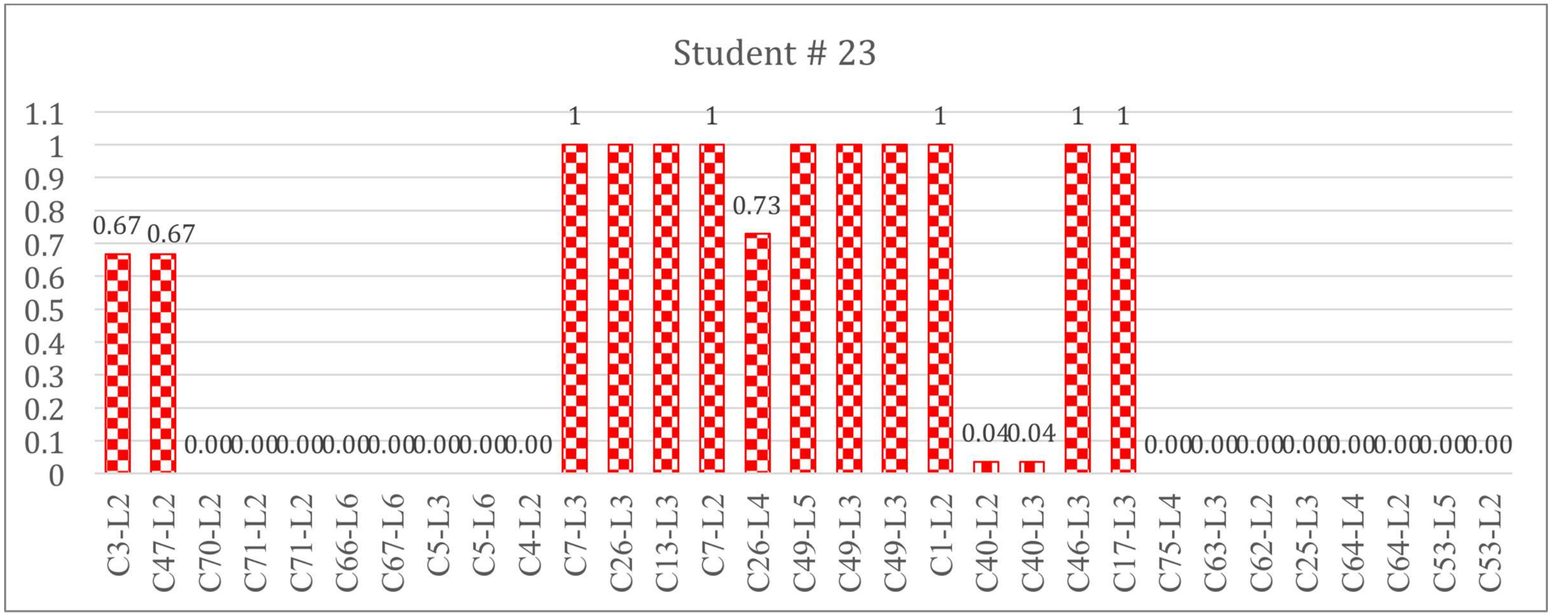
In conclusion, we observed that, if there are two correct answers to the questions asked about a concept at the same skill level, the result of the probability computation is equal to one, even though we consider the probability of error in each response. The probability of knowing concept c, that is evaluated based upon two questions which have conflicted responses, shows that the probability of knowing the concept is changing based on the type of questions. In order to determine the truth estimation, we suggest providing an error value to the responses based on the type of the question. If the question type is multiple choice, then the probability of the error is increased. Also, a higher probability of error is given to the question indirectly asking about the target skill, rather than the question directly asking about the skill. This is proved by the result of the computation of knowing the concepts in the real test. Table 6.3 illustrates the computation of the probability of knowing a concept, which is an element in DS or PS, by using the suggested Bayes' Formula and error values. Also, Figures 6.4 and 6.5 show the probability of knowing the concepts in the domain of DS according to the evaluation result of the perfect learner and the randomly selected learner (who didn't answer all the questions correctly), respectively. Whereas, Figures 6.6 and 6.7 show the probability of knowing the concepts in the domain of PS, according to the evaluation result of the perfect learner and the randomly selected learner, respectively. Also, Figure 6.8 and Figure 6.9 show the probability of 5 learners including the perfect learner and another 4 learners randomly selected from the domain of DS and PS respectively.



X : Concept # at certain skill level

Y : The probability of knowing the concept

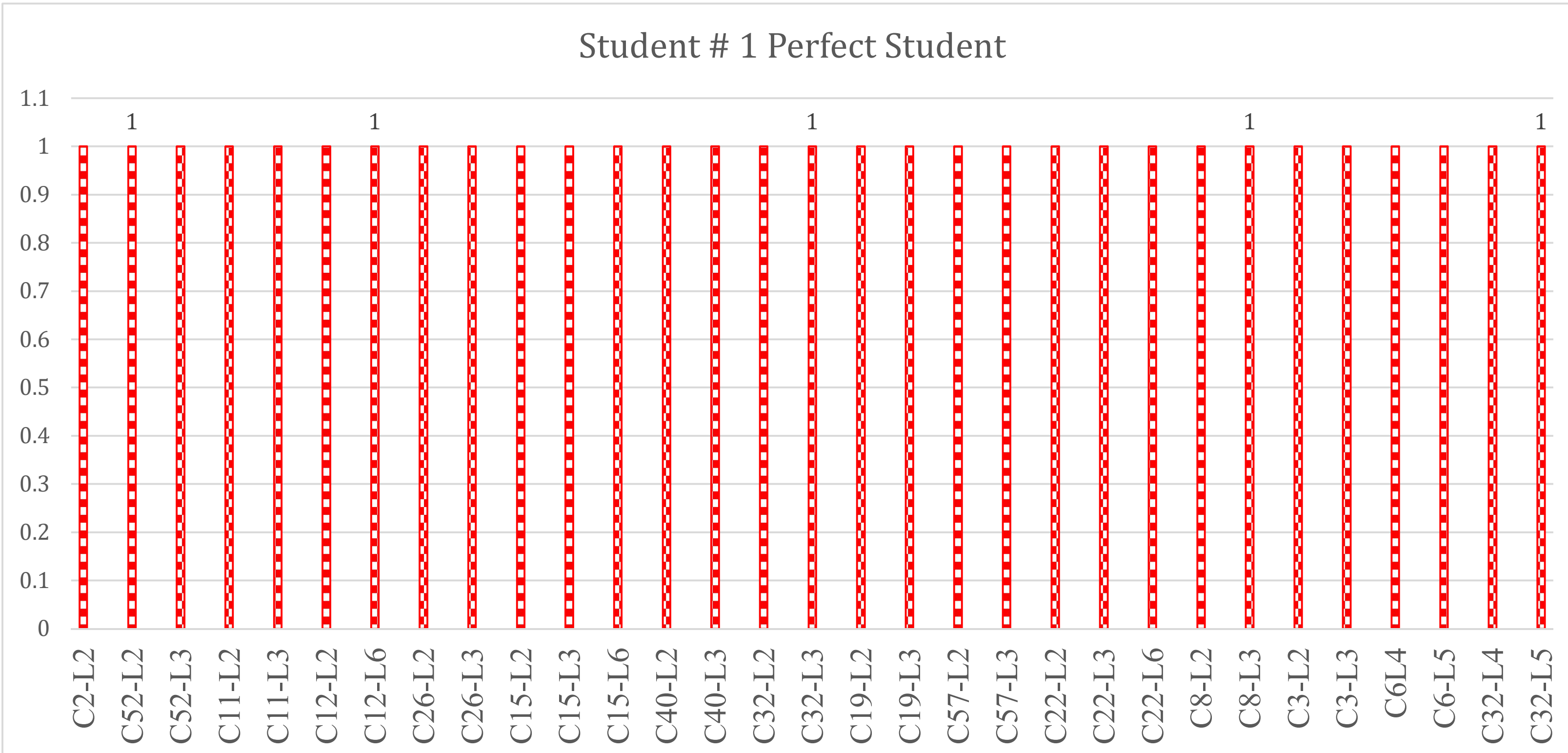
Figure 6.4 Probability of Knowing the Concepts in DS According to the Evaluation of the Perfect Learner



X: Concept # at certain skill level

Y: The probability of knowing the concept

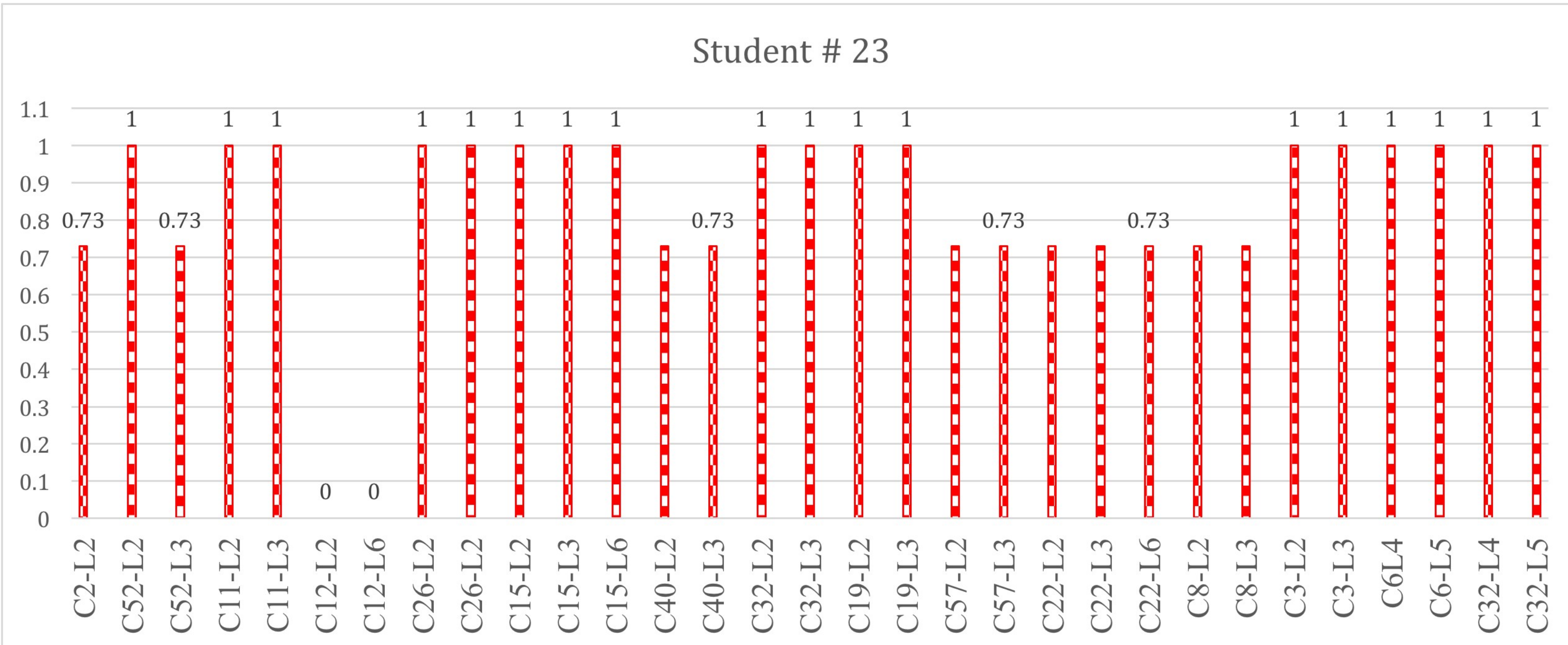
Figure 6.5 Probability of Knowing the Concepts in DS According to the Evaluation of Randomly Selected Learner



X: Concept # at certain skill level

Y : The probability of knowing the concept

Figure 6.6 Probability of Knowing the Concepts in PS According to the Evaluation of the Perfect Learner



X: Concept # at certain skill

Y: The probability of knowing the concept

Figure 6.7 Probability of Knowing the Concepts in PS According to the Evaluation of Randomly Selected Learner

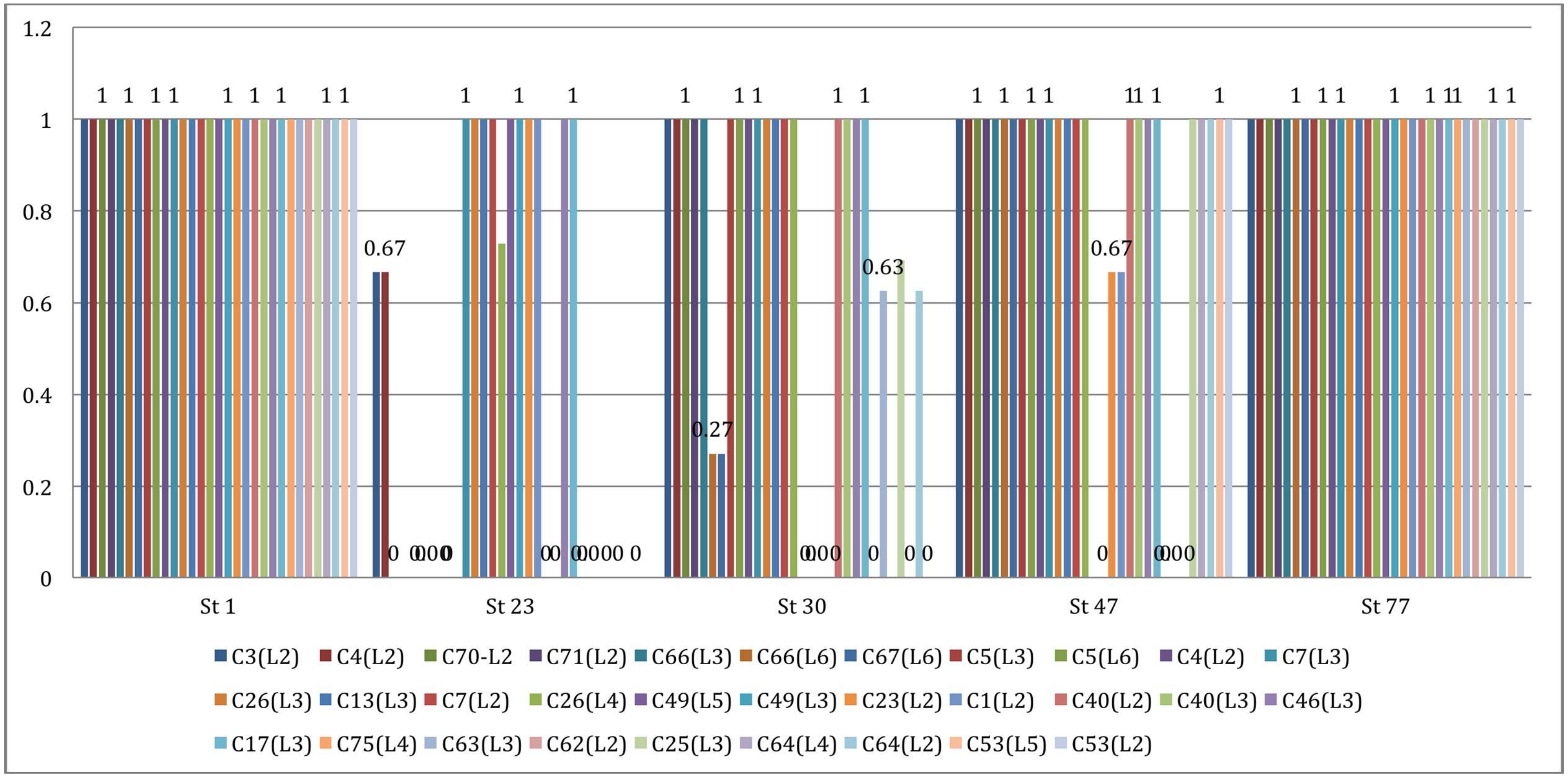
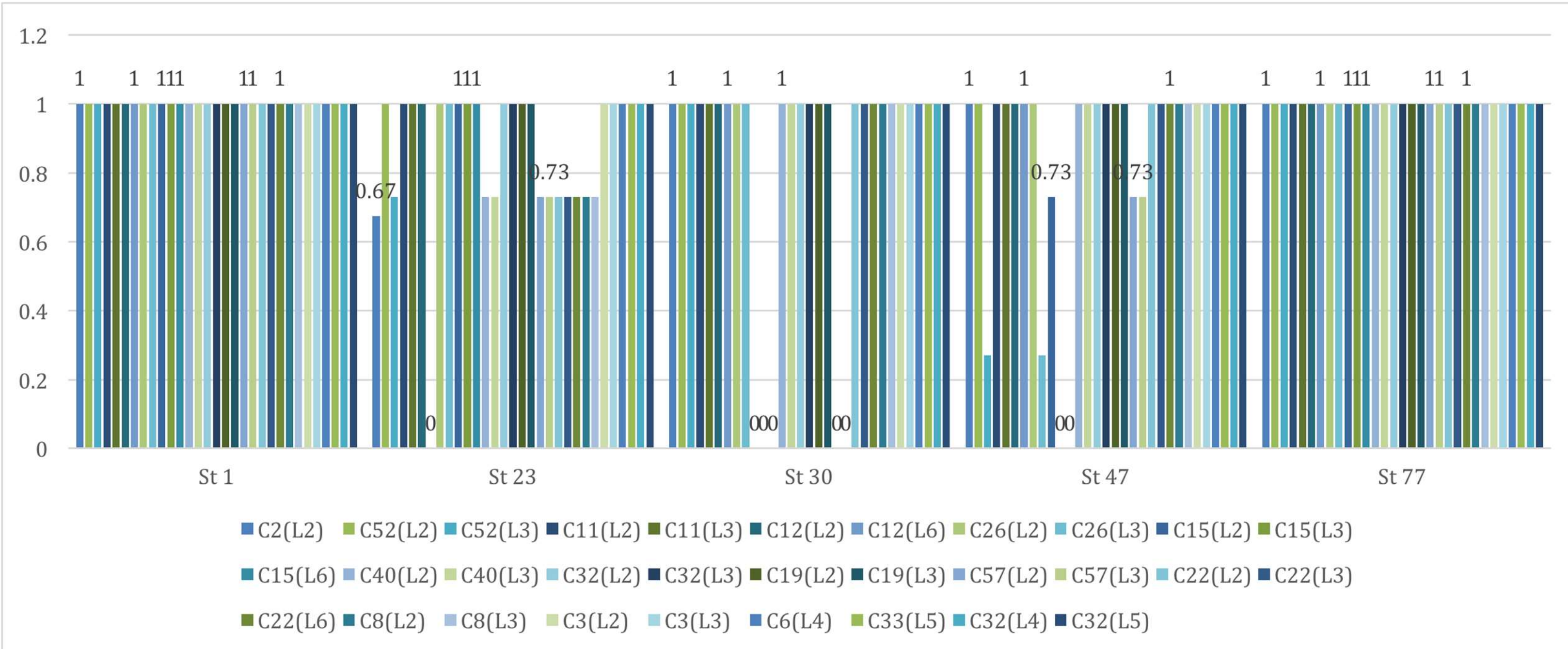


Figure 6.8 The Probability of Knowing the Concepts in DS

X: Learner Number
Y: Probability of knowing the concept

Example of 5 Learners Selected Randomly Including the Perfect Learner



X: Learner Number
Y: Probability of knowing the concept

Figure 6.9 The Probability of Knowing the Concepts in PS

Example of 5 Learners Randomly Selected Including the Perfect Learner

6.3 The Computation of the Probability of Knowing the Learning Object in the Domain VKS and VNS

In this calculation, we used the probability values of each concept based on the result in a previous section. I estimated that the tested concepts set of VS form the entire domain of the most advanced learning object and consequently forms the prerequisite set of the higher advanced learning object. Also, the estimated concepts set in DS indicates the probability of knowing the concepts of the lower advanced domain of the learning object. For the PS, it is logical to say that the estimated concepts set of PS indicates the probability of knowing the untested learning object at the same complexity level of either VS or DS. Thus, the probability of knowing the concepts set of VS indicates the probability of knowing the learning object. The probability of VKS indicates the probability of knowing the learning object, while the probability of VNS indicates the probability of not knowing the learning object. At the end of the calculation of each concept in VS, we finalized the study by three kinds of VKS: 1) The set of VKS with probability value equal to 1. 2) The set of VKS with probability equal to 0.5. 3). The set of VKS with probability value equal to 0.

Also, we finalize with two kinds of VNS: 1) The set of VNS with probability value equal to 1, and 2) The set of VNS with probability equal to 0.5. We took into account the concepts in the first probability value and the second probability value (1 & 0.5) to estimate the probability of knowing the learning object of VS.

Also, we took into account the concepts in the second and third type to estimate the probability of the verified known unknown learning object of VNS.

We used the concepts in the set of verified known with probability values equal to 1 and 0.5 to find out the probability of knowing the domain of the learning object, which involves the estimation of learning object concepts in the tested domain. The learning object is a concept in the domain related to all the dominant concepts in the Concept Space. In other words, it involves the learning objective. The computation of the probability of knowing and not knowing the learning object in the experiment domain is illustrated in Figure 6.10 section 6.4.

The equation used is the Bayes' Theorem (Equation 2). The difference here is in the calculation of the value of the probability of the related concepts to the learning object on condition of knowing the learning object $P(R|C_{LO})$, and the unconditional probability of knowing the Learning Object $P(C_{LO})$

$$P(C_{LO}|R) = \frac{P(\mathbf{R}|\mathbf{C_{LO}}) * P(C_{LO})}{P(\mathbf{R}|\mathbf{C_{LO}}) * P(C_{LO}) + P(\mathbf{R}|\bar{\mathbf{C}}_{\mathbf{LO}}) * P(\bar{C}_{LO})} \qquad \text{Equation 10,}$$

where

-C_{LO} denotes knowing the learning object concept.

- $P(C_{LO})$ is the unconditional probability of knowing the learning object concept. It is just the ratio of knowing concepts in the domain of the learning object to the entire concepts in the domain of the learning object.

$$P(C_{LO}) = \frac{\text{Total number of the related knowing concepts in VKS } \left(VKS_{C_{LO}}\right)}{\text{Total number of the related concepts in the entire domain of the Learning Object } \left(VS_{C_{LO}}\right)}$$

$= \frac{|\, VKS_{c_{LO}}|}{|VS_{c_{LO}}|}$ Equation 11

$VKS_{c_{LO}}$ is the set of the knowing concepts in VKS and related to the learning object.

$VS_{c_{LO}}$ is a set of the concepts in the domain of the learning object.

$P(\bar{C}_{L0})$ is the unconditional probability of not knowing the learning object concept, it is just the rate of not knowing concepts in the domain of the learning object to the entire concepts in the domain of the learning object.

$$P(\bar{C}_{L0}) = \frac{\text{Total number of the related not knowing concepts in VNS } \left(\mathrm{VNS}_{C_{LO}}\right)}{\text{Total number of the related concepts in the entire domain of the Learning Object } \left(\mathrm{VS}_{C_{LO}}\right)}$$

$= \frac{|\, \mathrm{VKS}_{\bar{c}_{LO}}|}{|\mathrm{VS}_{c_{LO}}|}$ Equation 12

$P(R|C_{LO})$ = The probability of knowing the concepts in the learning object domain for a single learner.

$= \sum_{i=1}^{n=|VKS|} P\,(C_i)$ Equation 13

C_i is a concept in the set of VKS

$P(C_i)$ is a probability of knowing a concept C_i

$P(R|C_{LO})$ is the conditional probability of knowing the concepts in the learning objects given knowing the learning object.

6.4 The Experimental Computation of the Probability of Learning Object in the Domain of VKS & VNS

The experimental result of the computation is the probability of knowing the learning object in VS. The result reveals the probability of knowing the learning object of

154 learners who participated in the test. This result is based on the tested concepts in the dataset VS. The concepts set of VS are the most advanced concepts in the learning domain. Figure 6.10 shows the probability of knowing and not knowing the concepts in the learning objects of VS domain. The Figure 6.10 shows the probability of knowing the learning object domain of 30 learners randomly selected from the 154 participants.

The Figure 6.10 shows the biography of the 30 learners. The red line illustrates the probability value of the not knowing learning object concepts. The blue line illustrates the probability values of the knowing learning object concepts. The probability distribution is between 0 to 1.

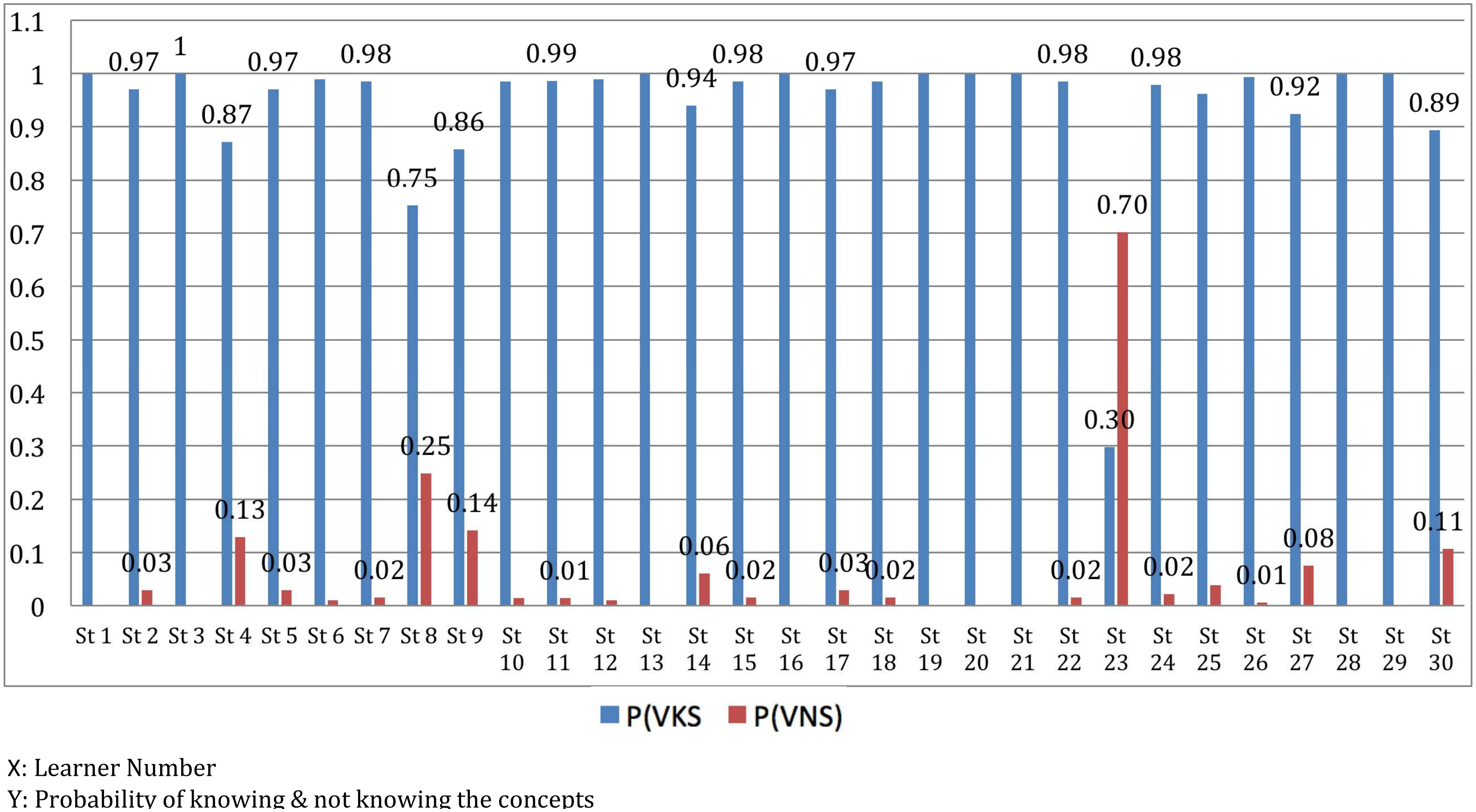
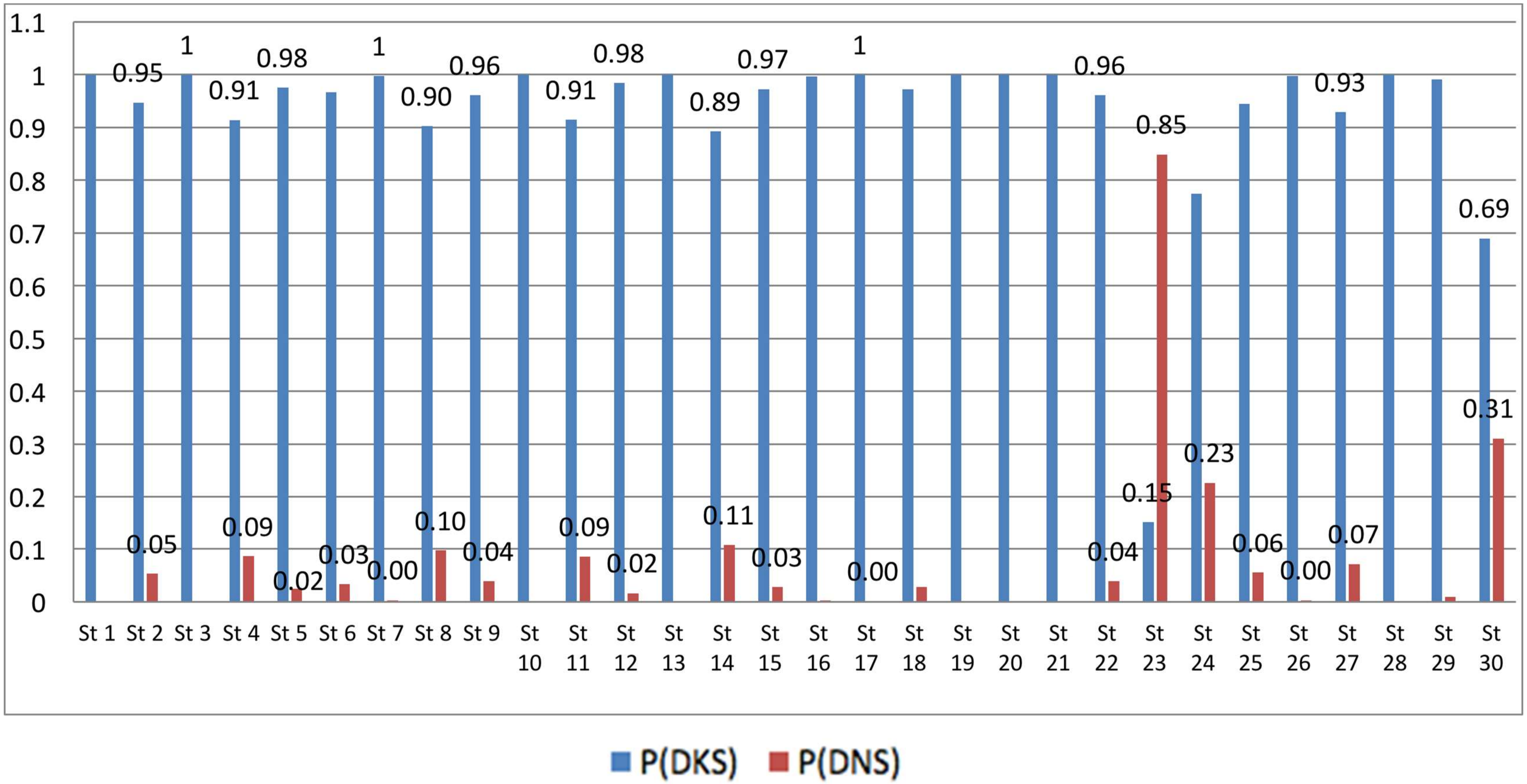


Figure 6.10 The Probability of Knowing and Not Knowing the Learning Objects of the 30 Learners in the VS Domain

6.5 The Experimental Computation of the Probability of Learning Object in the Domain of DKS and DNS

In this calculation, we use the result of the computation of the probability of each concept in the domain of DS which was illustrated in section 6.2. The tested concepts set of DS forms the entire domain of the prerequisite concept set of the learning object domain and consequently forms a part of the prerequisite of the prerequisite set of the learning object. Thus, the probability of knowing the concepts set of DS indicates the probability of knowing the prerequisite set of the lower advanced concepts of the learning object. The probability of DKS indicates the probability of the knowing concepts in the prerequisite set of the domain of learning object. The probability of knowing the domain in DNS indicates the probability of the not knowing concepts in the prerequisite set of the domain of the learning object. Figure 6.11 illustrates the probability of knowing and not knowing the learning objects in the DS domain. The blue line indicates the probability of knowing the learning object concepts of the related learner in the domain of DKS, whereas the red line indicates the probability of not knowing the learning object of the related learner in the domain.

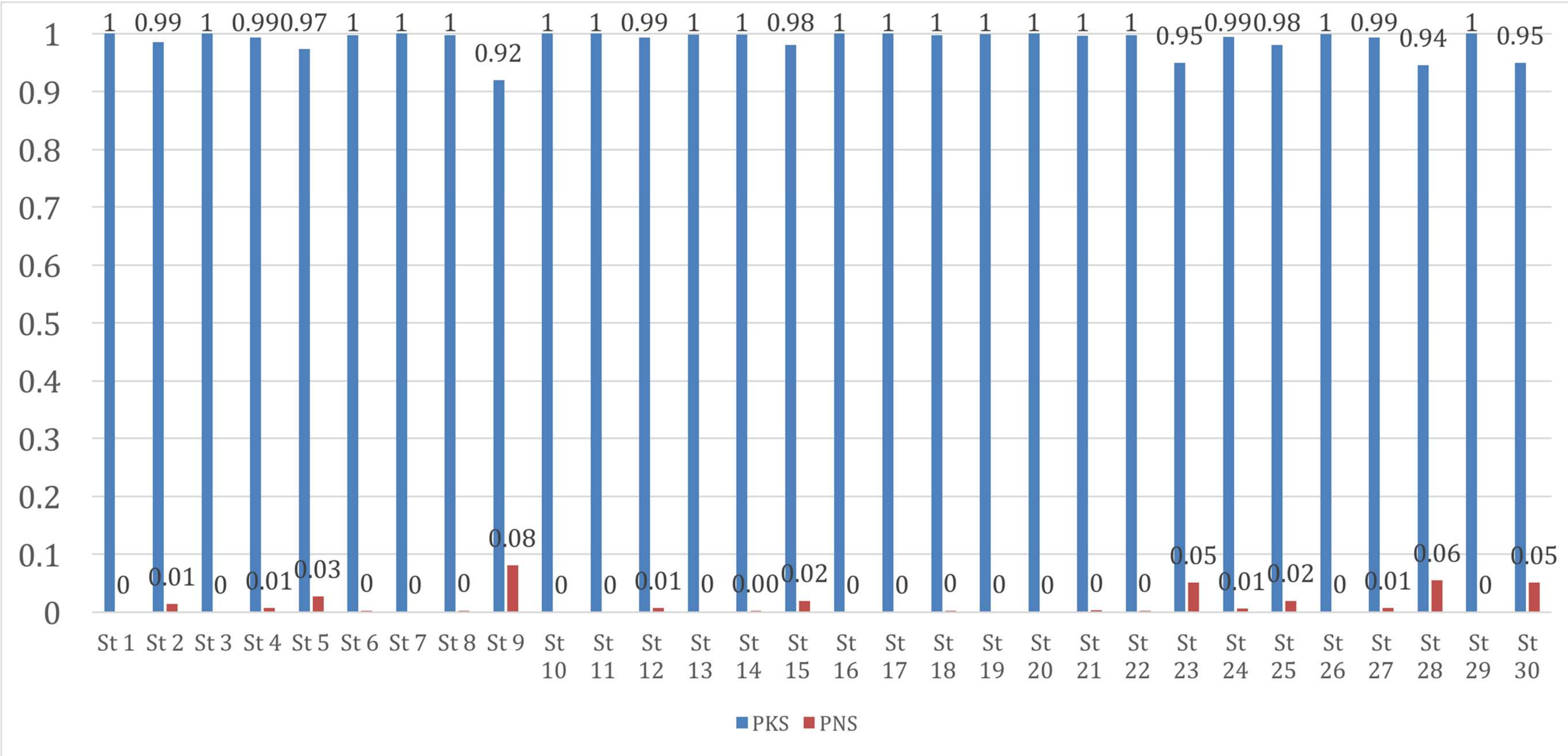


X: Learner Number
Y: Probability of knowing & not knowing the concepts

Figure 6.11 The Probability of Knowing and Not Knowing the Learning Objects in the DS Domain of 30 Learners

6.6 The Experimental Computation of the Learning Object in the Domain of PKS and PNS

In this calculation, we used the result of the computation of the probability of each concept in the domain of PS in a previous section. The tested concepts set of PS forms the entire domain of the prerequisite concept set of the learning object domain and, consequently, they form the prerequisite set of the learning object. Thus, the probability of knowing the concepts set of PS indicates the probability of knowing the prerequisite set of the lower advanced concepts of the learning object. The probability of PKS indicates the probability of the knowing concepts in the prerequisite set of the domain of learning object. The probability of PNS also indicates the probability of the not knowing concepts in the prerequisite set of the domain of the learning object. Figure 6.12 illustrates the probability of knowing and not knowing the learning objects in the PS domain. The blue line indicates the probability of knowing the concepts of the learning object of the related learner in the domain of PKS, the red line indicates the probability of not knowing the concepts of the learning object of the related learner in the domain of PNS.



X: Learner Number
Y: Probability of knowing & not knowing the concepts

Figure 6.12 Probability of Knowing and Not Knowing the Learning Objects in the PS Domain of 30 Learners.

CHAPTER 7

The Experiments to Validate the Concept States of the TCS^2 theory

This chapter provides two experiments aimed at validating the Concept States proposed in the assessment. Also, the experiments measure the accuracy and the efficiency of the methodology. The accuracy is proved by the comparison between the estimated probability of knowing the concepts in DS and PS, and the true probability of the concepts in DS and PS by real responses. The efficient is proved by analyzing the size of the footprint of the perfect learner. Also, in this chapter, we explain some studied applications of the proposed TCS^2 theory.

7.1 The Experiments to Validate the Proposed Concept States

We organized two experiments aimed at validating the Concept States proposed in the assessment. Each experiment was organized to prove the precise estimation of the proposed three sets of Concept States: VKS, DKS, and PKS. Also, we proved accuracy and the efficiency of the TCS^2 theory to maximize the estimation of measurement the concepts from few tested concepts. In this setup, the questions are specially designed to directly test a certain skill level of each concept belonging to the concept set. We call these questions Direct Questions (DQ) and the directly tested concept skills as Direct Concept Mapped Skills (DCMS). The DQ tests the identical skill level, which has been tested by the instructor, but it directly specifies the level of the concept. We call the normal questions, prepared by the instructor, as Open Questions (OQ). The open question is any question prepared by the instructor and which could implicitly test the skill level. When the open question is analyzed, then we can conclude which skill level was included in that open question. Therefore, two types of questions are offered: (1) OQ. (2) DQ.

To detect the DCMS, the TCS^2 Assessment Analytics are applied to each question. The DQ, then, is prepared to directly test the DCMS. Subsequently, for each detected concept C_X at a certain level k, either Verified or Derived or Potential, DQ are designed for directly verifying the matching of the related skills between OQ and DQ, based on the relation within the three sets of Concept States of VKS, DKS, and PKS. We introduced a test for validating the relation of the TCS^2 theory. The match of the learner knowledge between the two types of questions was calculated. We found out the match of the correct answers between the associated skills. We gave a value of 1 to each correct answer for each tested skill by OQ and DQ. If the answer was wrong, the tested skills were given a value of 0. Thus, if the learner's answer to the identical tested skill had the same value either 0 or 1 in DQ and OQ then the matching value will be 1, otherwise it will be 0. Table 7.1 illustrates the matching of logical computation.

Table 7.1 Computation of the Matching Between the OQ & DQ

Value of Answer to Open Question	Value of Answer to Direct Question	Matching Result	Validating Theory
1	1	1	Correct (+)
0	0	1	Correct (−)
1	0	0	False (+)
0	1	0	False (−)

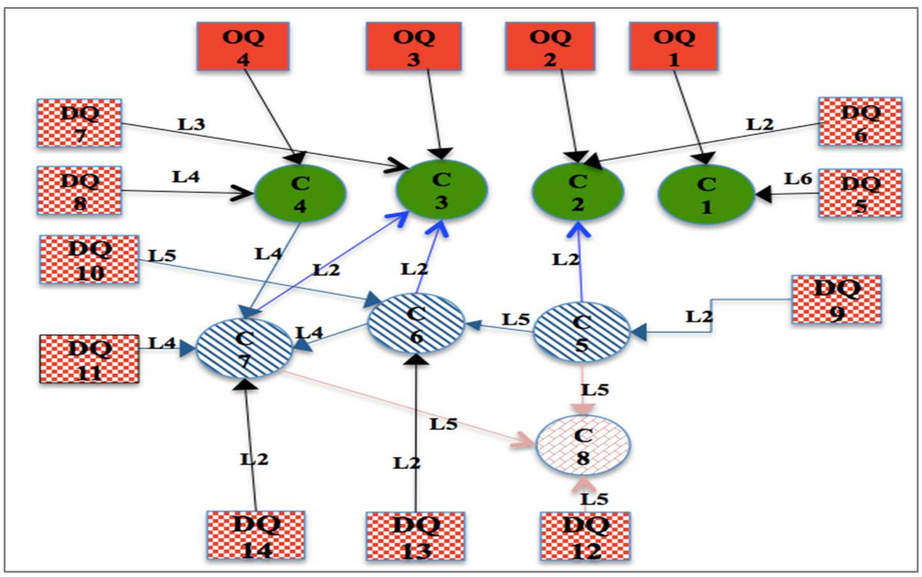


Figure 7.1 Illustration of Skills Counting

The matching percentage of the learner's knowledge is observed at each level, i.e.: skill level k =2 which is the Understand level, skill level k =3 which is the Apply level, skill level k = 4 which is analyze level, skill level k = 5 which is the Evaluate level, and skill level k = 6 which is the Create level; to verify the qualification of the Concept States of the three sets: VKS, DKS, and PKS. The tested skills are calculated by counting each tested level of each concept in the test. Figure 7.1, and Tables 7.2 and 7.3 give an illustration of an example

Table 7.2 Counting Skills Result According to Figure 7.1

Type of Questions	Type of Estimated Skill	Bloom Link	Concepts Counting
OQ & DQ	V	2	1
OQ & DQ	V	3	1
OQ & DQ	V	4	1
OQ & DQ	V	6	1
DQ	D	2	3
DQ	D	3	1
DQ	D	4	1
DQ	D	5	1
DQ	P	5	1

Table 7.3 The Related Questions that have to be Matched Based on Figure 7.1

Relation Type	C#	Level#	Related OQ	Related DQ
V	1	6	1	5
V	2	2	2	6
V	3	3	3	7
V	4	4	4	8
D	5	2	2	9
D	6	2	3	13
D	7	2	3	14
D	6	5	3	10
D	7	4	3	11
D	6	5		9 & 10
D	7	4		11 & 13
P	8	5		9 &12
P	8	5		11&12

Accordingly, to validate the proposed theories, we conducted two separate experiments based on the matching calculation method.

7.2 The First Experiment

7.2.1 The Validation Test Setup

The test is introduced online in one session. The participants are 45 graduate learners, who attend an algorithm class in the Computer Science department: CS 46101 Design and Analysis of Algorithms. Two types of questions, DQ and OQ, are given to the learners. OQ are typical to the questions that were already given to the learners in two successive tests prepared by the instructor. We called the first test introduced by the instructor OQ_1 and the second OQ_2. We combined them with our prepared questions DQ and gave an online test in one session to the participants at the end of the semester. Accordingly, DQ are classified as DQ_1 and DQ_2, which are respectively based on OQ_1 and OQ_2. The online test contains 66 questions. The first 9 questions are selected from OQ_1 and OQ_2, and these form OQ. The remaining 57 questions are selected from DQ_1 and DQ2, and these form DQ.

7.2.2 The Accuracy Result

The result analysis of VS is illustrated in Figure 7.2. As is evident, the correct result shows the highest accuracy at all levels. Level 4 and level 6 show the highest accuracy between all levels with accuracy of 96%, where in level 4 the correct (+) has 89% of matching results and the correct (−) has 7% of matching result whereas, level 6 has correct (+) accuracy 86%, and correct (−) accuracy 10%. Second highest accuracy is level 3 with 92%, then level 2 and level 5 with achievements 89%, and 82% respectively. Similarly, for determining the accuracy of DS, we observed the matching of the learners' answers between the related skills based on DS method. The DS analysis result is illustrated in Figure 7.3. As is evident, the correct result shows the highest accuracy at all levels. Level 6 shows the highest correct accuracy: 92%. The second highest matching percentages are levels 3 and 5 with the same accuracy: 90%, then level 2 and level 4 with scores of 85% and 86%, respectively. The PS analysis result is illustrated in Figure 7.4. As is evident, the correct result shows the highest accuracy at all levels. Level 5 shows the highest accuracy: 89% between all levels. The second highest positive accuracy is level 6 with 86%, then level 3, level 4, and level 2 with scores of 83%, 82%, and 78%, respectively. As observed in Figure 7.4, the false (−) which, is the result of matching 0-1 and 0-0, is ignored since it is an inappropriate matching to assess PS relation. We cannot estimate what the learner will be ready to know if his answer to the question of the related concept is false.

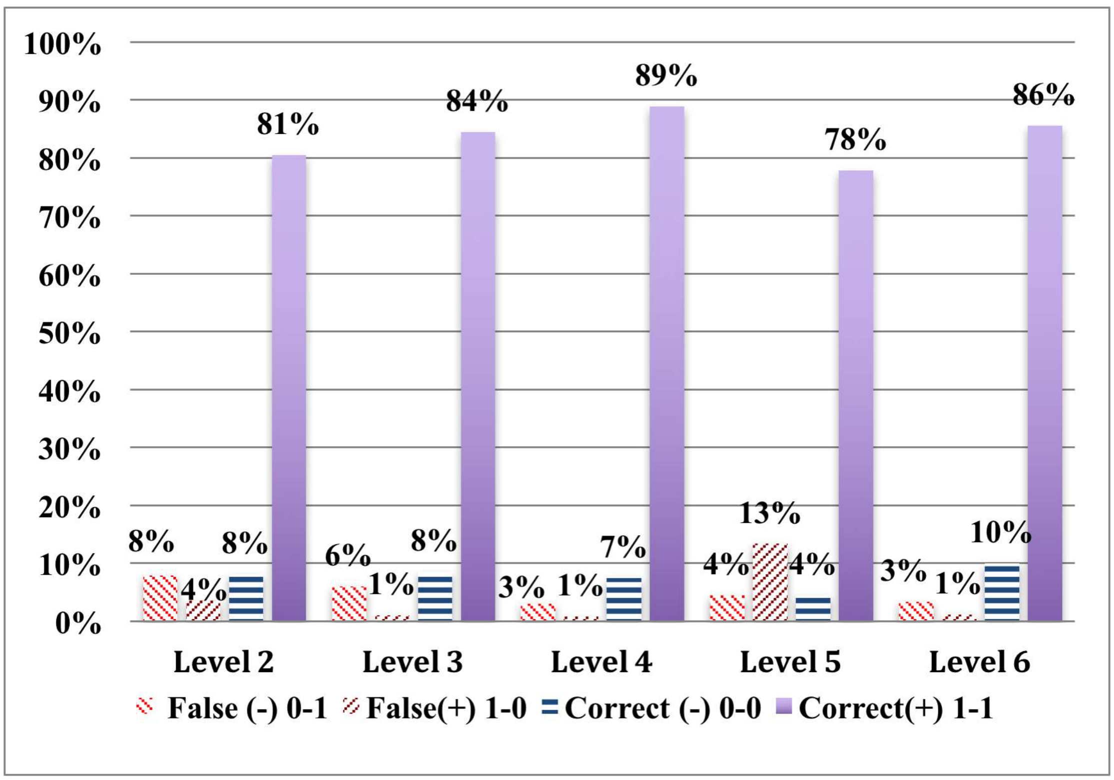


Figure 7.2 The Percentage of the Matching of Skills in the Set of VS

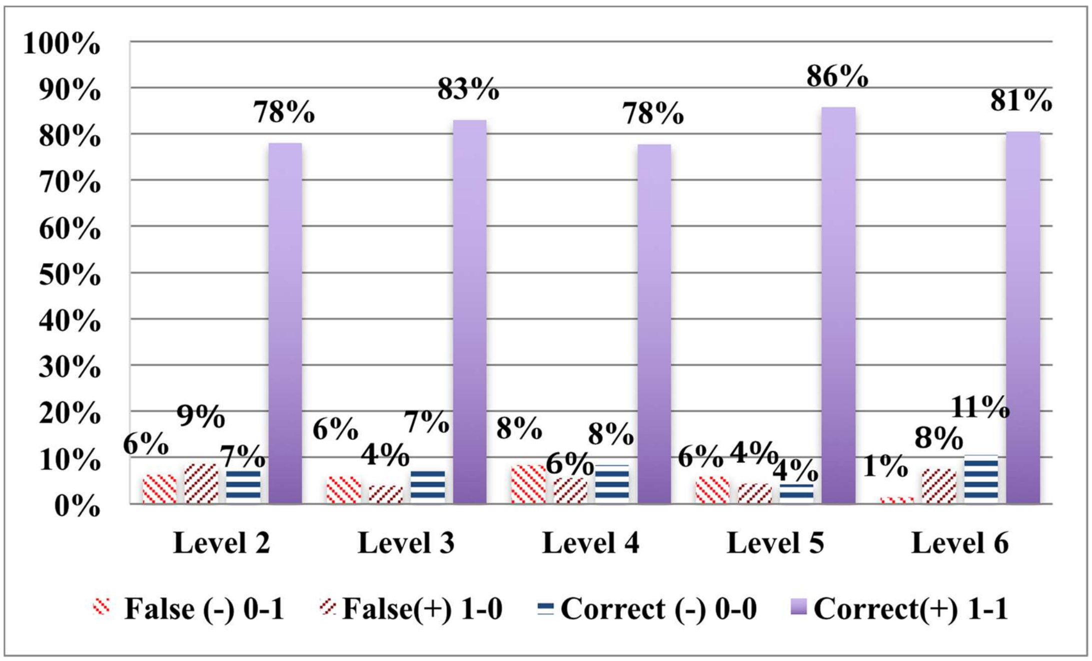


Figure 7.3 The Percentage of the Matching of Skills in the Set of DS

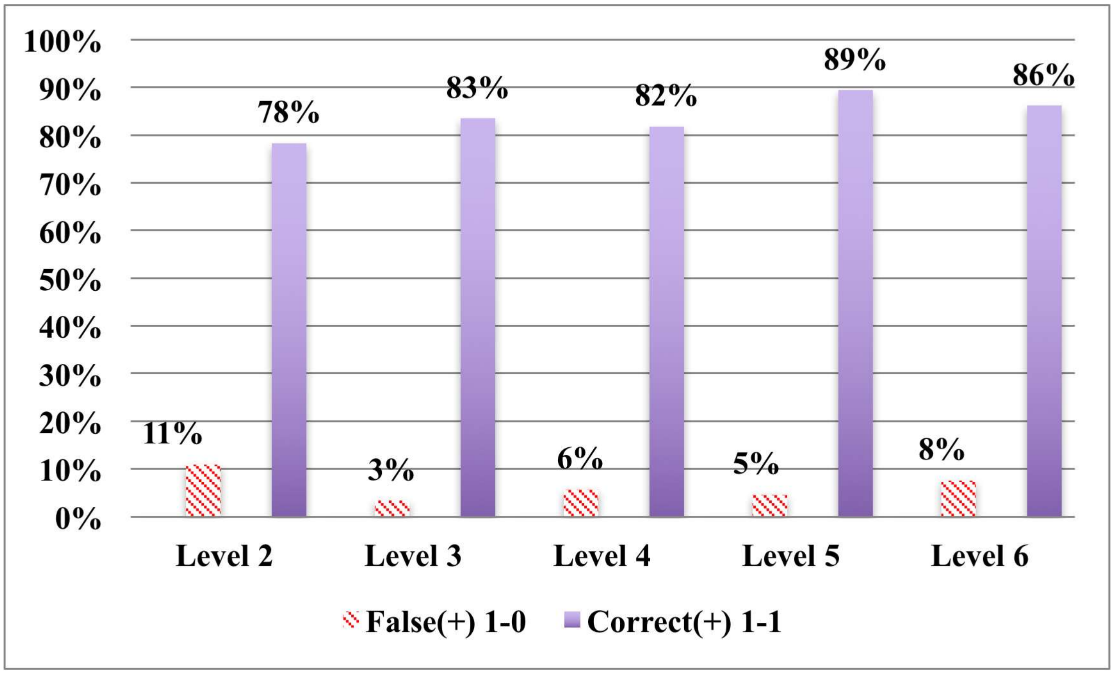


Figure 7.4 The Percentage of the Matching of Skills in the Set of PS

7.2.3 The Size of Footprint

The experiment proves that using the proposed methods of the Concept States optimize the knowledge assessment. The result of the evaluation of the perfect learners shows that the amount of the estimated knowledge of the assessed learner could be increased by at least 3 times over the conventional assessment, which uses just numerical methods. This is because the tested set of concepts in DQ is prepared based on the OQ and directly tests the estimated learner knowledge. If the learner answered a question directly testing a certain level of skills, then we can measure his estimated knowledge of the associated skills without testing them. Suppose a directly tested set by OQ is prepared by the instructor and it is cited in a question OQ = [a, b, c, d, e]. The assessment analytic realizes the level of each tested concept, and the set will become OQ = [$VKS(a)_2$, $VKS(b)_3$, $VKS(c)_4$, $VKS(d)_5$, $VKS(e)_6$], where, $VKS(a)_2 \in$ level 2 of RBT. $VKS(b)_3 \in$ level 3 of RBT. $VKS(c)_4 \in$ level 4 of RBT. $VKS(d)_5 \in$ level 5 of RBT. $VKS(e)_6 \in$ level 6. Similarly, the estimated set of DS and PKS could be realized from the basic tested set. It is clear that the basic tested concepts are a, b, c, d and e, and the counting number of tested skills is 6. We call the assessment analytic process EVAL(OQ); the estimated skills set at each level of VKS, DKS and PKS, is the experiment footprint; and the counting number of these skills, is the size of the footprint. If the learner correctly answered the entire OQ, the learner will be considered the perfect learner and the experiment optimal answers is his answer. The perfect learner answers are used to calculate the experiment footprint and the size of footprint of each Concept State. Table 7.4 and Figure 6.5 show the size of footprint according to the perfect learner. As is evident, the size of VKS footprint is 25, the size of DKS footprint is 23 and the size of PKS footprint is 11, which means that if the learner answers the instructor OQ correctly, we can tell he knows certain levels of the 48 concepts, and he is ready to know additional 11 concepts also in certain levels of each of them, even though he was tested particularly for 25 concepts. By combining the presented knowledge assessment theories with cognitive relation, we can maximize the amount of the information estimated about the knowledge of the assessed learner. However, the benefit of adding the cognitive level as measurement parameter was studied to evaluate real assessment introduced by the instructor. This issue is studied in the section 7.4, which is on the application of the proposed TCS^2 theory.

Table 7.4 Size of Footprint

Skill Level parameter	Verified	Derived	Potential
L2	10	10	1
L3	9	6	3
L4	3	3	1
L5	1	2	3
L6	2	2	3
Sum	25	23	11

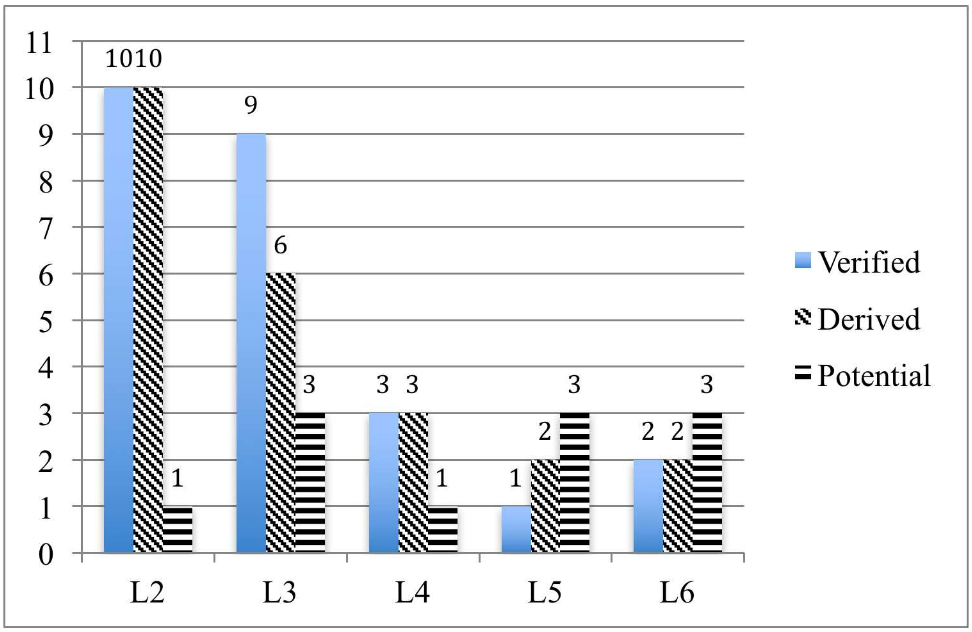


Figure 7.5 The Size of Footprint

7.3 The Second Experiment to Validate the Accuracy of the Concept States

7.3.1 The Validation Test Setup

The test is introduced online in one session. The participants are 154 graduate learners, attending the CS 61002 Algorithms and Programming class in the Computer Science department. The two types of questions (DQ and OQ) were administered to the learners. The OQ are typical of the kind of questions that are usually given to the learners for their midterm exam. We prepared the DQ type of questions from the midterm questions and gave an online test in one session to the participants at the end of the semester. The online test contained 47 questions. The first 9 questions were selected from the OQ, and the remaining 38 questions were direct questions asking about the exact skills that were extracted from the analysis of the OQ.

7.3.2 The Accuracy Result

The result analysis of VS is illustrated in Figure 7.6. As is evident, the correct result shows the highest accuracy at all levels. Levels 2, 3 and 6 show the highest accuracy between all levels with the same accuracy: 96%, where level 2 has correct positive accuracy of 84% and correct negative accuracy of 12% of matching results, level 3 has correct positive accuracy of 83% and negative accuracy of 13% of matching results, and level 6 has correct positive accuracy of 83% and correct negative accuracy of 13% of matching results. The second highest accuracy between all levels is 95% of level 4, with correct positive accuracy of 90% and correct negative accuracy of 5% of matching results. The last highest accuracy is level 5 with accuracy of 89% since the correct (+) has

58% of matching results and the correct (−) has 31% of matching results. Similarly, for determining the accuracy of DS, we observed the matching of the learners' answers between the related skills based on DS method. The DS analysis result is illustrated in Figure 7.7. As it is evident, the correct result shows the highest accuracy at all levels. Level 6 and 5 show the highest correct accuracy: 98%. Level 6 has correct (+) accuracy of 95% of matching results and has correct (−) accuracy of 3% of matching results, and level 5 has correct (+) accuracy of 67% of matching results and has correct (−) accuracy of 31% of matching results. The second highest matching percentage are levels 2 and 3 with the same accuracy: 96%. Level 2 has correct (+) 81% and correct (−) 15% of matching results, and level 3 has correct (+) accuracy of 84% and correct (−) accuracy of 12% of matching results. The third highest matching percentage is level 4 with accuracy of 94%, since the correct (+) has 71% of matching results and the correct (−) has 23% of matching results. The PS analysis result is illustrated in Figure 7.8. As is evident, the correct result shows the highest accuracy at all levels. Level 4 shows the highest accuracy, 92%, between all levels. The second highest correct accuracy is level 2 with 89%, then level 3, level 5, and level 6 with scores of 86%, 81%, and 79% respectively. As observed in Figure 7.8, the false (−) with 0-1 and 0-0 matching is ignored, since it is an inappropriate matching to assess PS relation. We cannot estimate what the learner will be ready to know, if his answer to the question of the related concept is false.

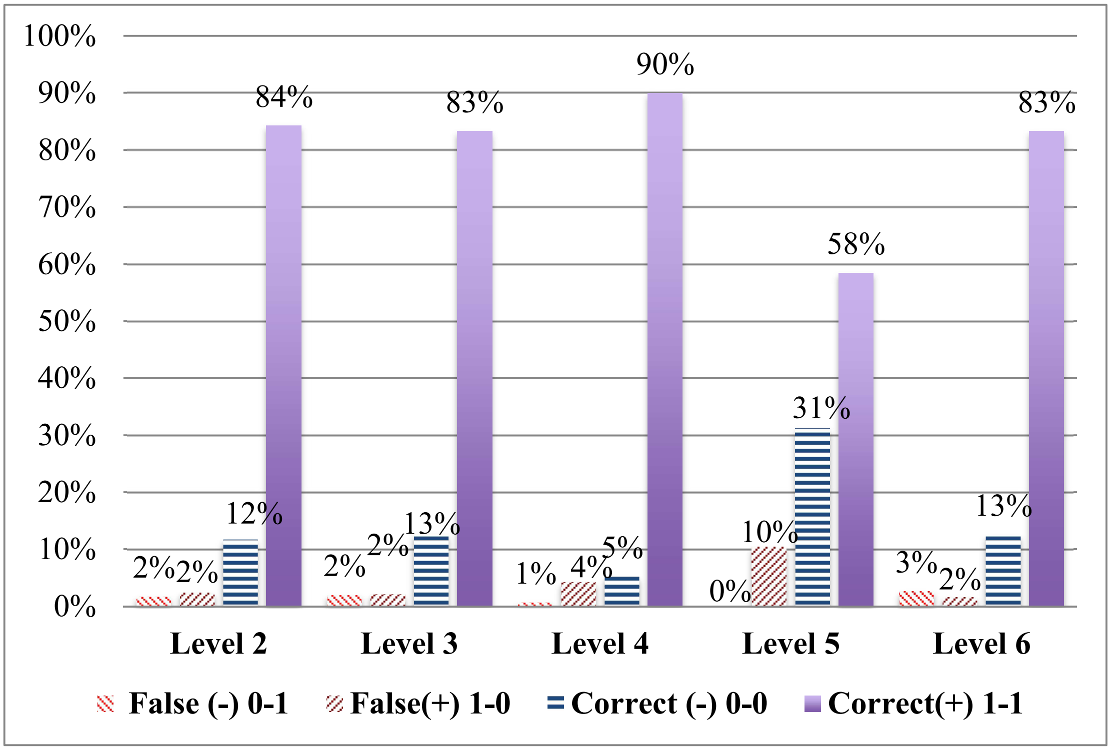


Figure 7.6 The Percentage of the Matching Skills in the Set of VS

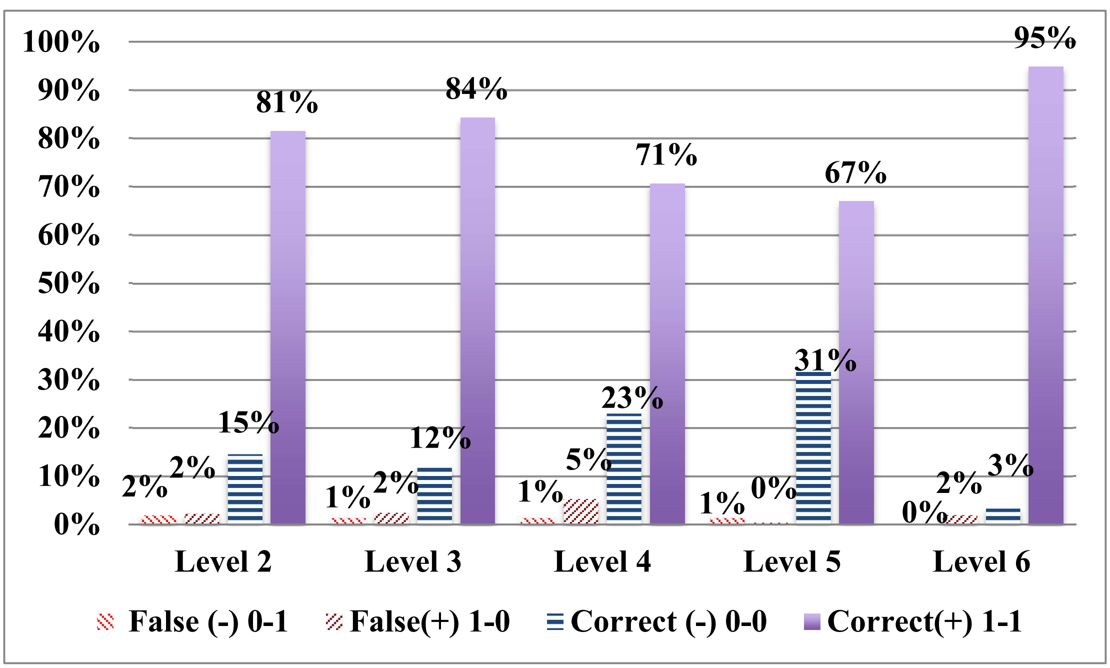


Figure 7.7 The Percentage of the Matching Skills in the Set of DS

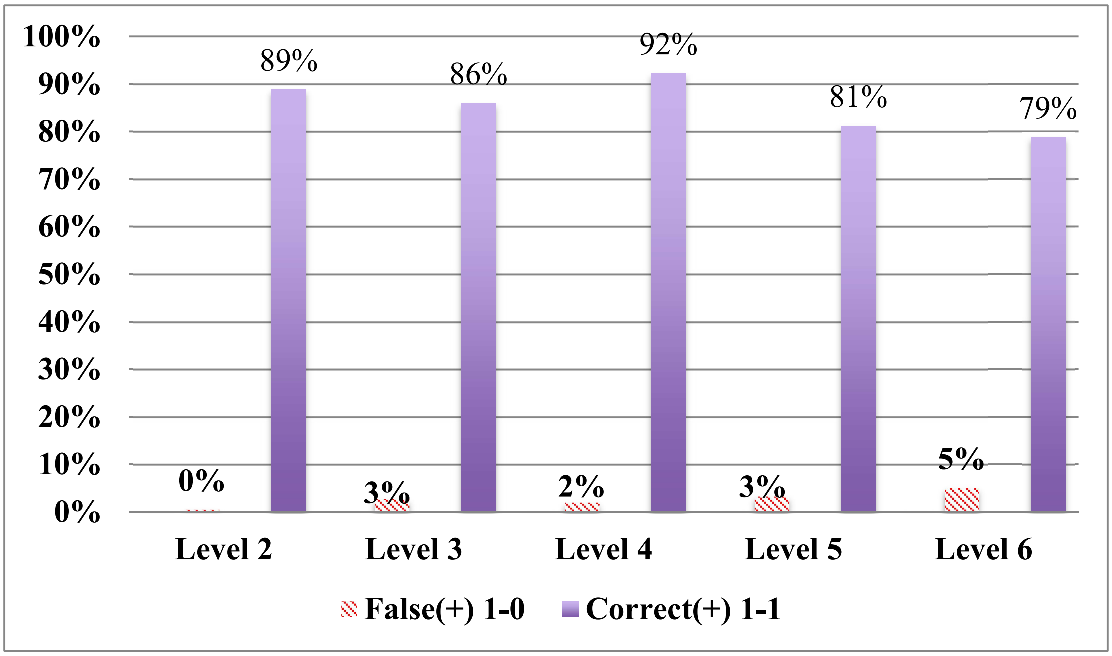


Figure 7.8 The Percentage of the Matching of Skills in the Set of PS

7.3.3 The Size of Footprint

As well as in the result of experiment 1, this experiment proves that using the proposed methods of the Concept States optimizes the knowledge assessment. The result of the evaluation of the perfect learner shows that the amount of the estimated knowledge of the assessed learner could be increased by at least 3 times over the conventional assessment which uses just numerical methods. The perfect learner's apostrophe answers are used to calculate the experiment footprint and the size of footprint of each Concept State. Table 7.5 and Figure 7.9 show the size of footprint according to the perfect learner. As evident the size of VKS footprint is 18, the size of DKS footprint is

31 and the size of PKS footprint is 31, which means that if the learner answered the instructor OQ correctly, then we can tell he knows certain levels of each of these 18 concepts, and he is ready to know an additional 31 concepts in certain levels of each of them, even though he was tested particularly for 49 skills. By combining the presented knowledge assessment theories with cognitive relation, we can maximize the amount of the estimation knowledge of the assessed learners.

Table 7.5 Size of Footprint

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

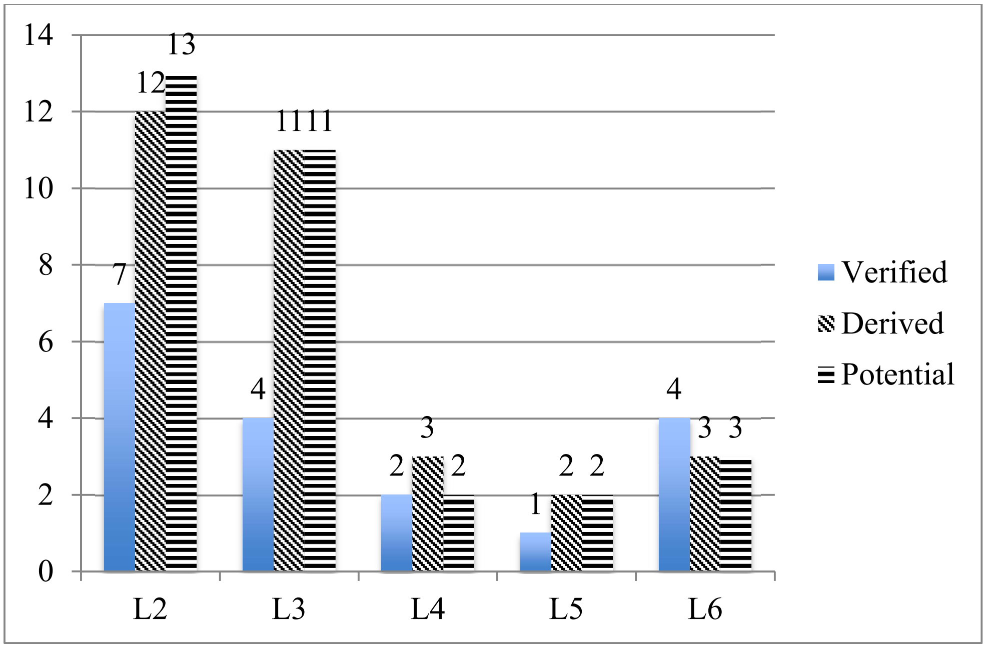


Figure 7.9 The Size of Footprint

However, the benefit of adding the cognitive level as measurement parameter is studied to evaluate real assessment introduced by the instructor. This issue is studied in section 7.5.

7.4 The Accuracy of the Estimated Probability of Knowing the Concept Based on the Methods.

We proved the accuracy of the estimated knowledge based on the proposed methods DS and PS by showing the comparison between the estimated probability of knowing the concepts based on the methods and the real probability of knowing the concepts based on the real response to the question asked about the certain concept at the certain estimated skill. We use the information of Experiment 2 to prove the accuracy of the methods. In addition, we take advantage of the existing information of the estimated knowledge and the real knowledge to show the accuracy of the probability of knowing the concepts. The structure of the concepts relation based on the introduced exam at the Experiment 2 is illustrated in the Figure 7.10. 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. Figures 7.11, 7.12 show the probabilities of knowing the concepts of the perfect learner and randomly selected learner (who didn't answer all the questions correctly), respectively.

Similarly, we make a comparison of the probability of knowing the concepts based on PS. Figures 7.13 and 7.14 show the accuracy of the probabilities according to the estimation of PS method. Figure 7.13 shows the comparison of the evaluation of the perfect learner, whereas Figure 7.14 shows the comparison of the evaluation of randomly selected learner.

In the Figures 7.11, 7.12, 7.13 and 7.14, 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.

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 our investigation, we increase the probability of errors values in the probability of knowing the estimated concept. In the direct question which is asked about the same estimated concept the errors values are less. Also, we assign an error value to the type of the question, whether it is multiple choice or an open-ended question. 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 askes directly about the concept. From this point, we observed little difference 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, OQ and DQ. The second question of DQ is to confirm the first response to the question of OQ type. Figures 7.15 and 7.16 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. We assigned the same probability of the error value to both types of the questions, OQ and DQ. 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.

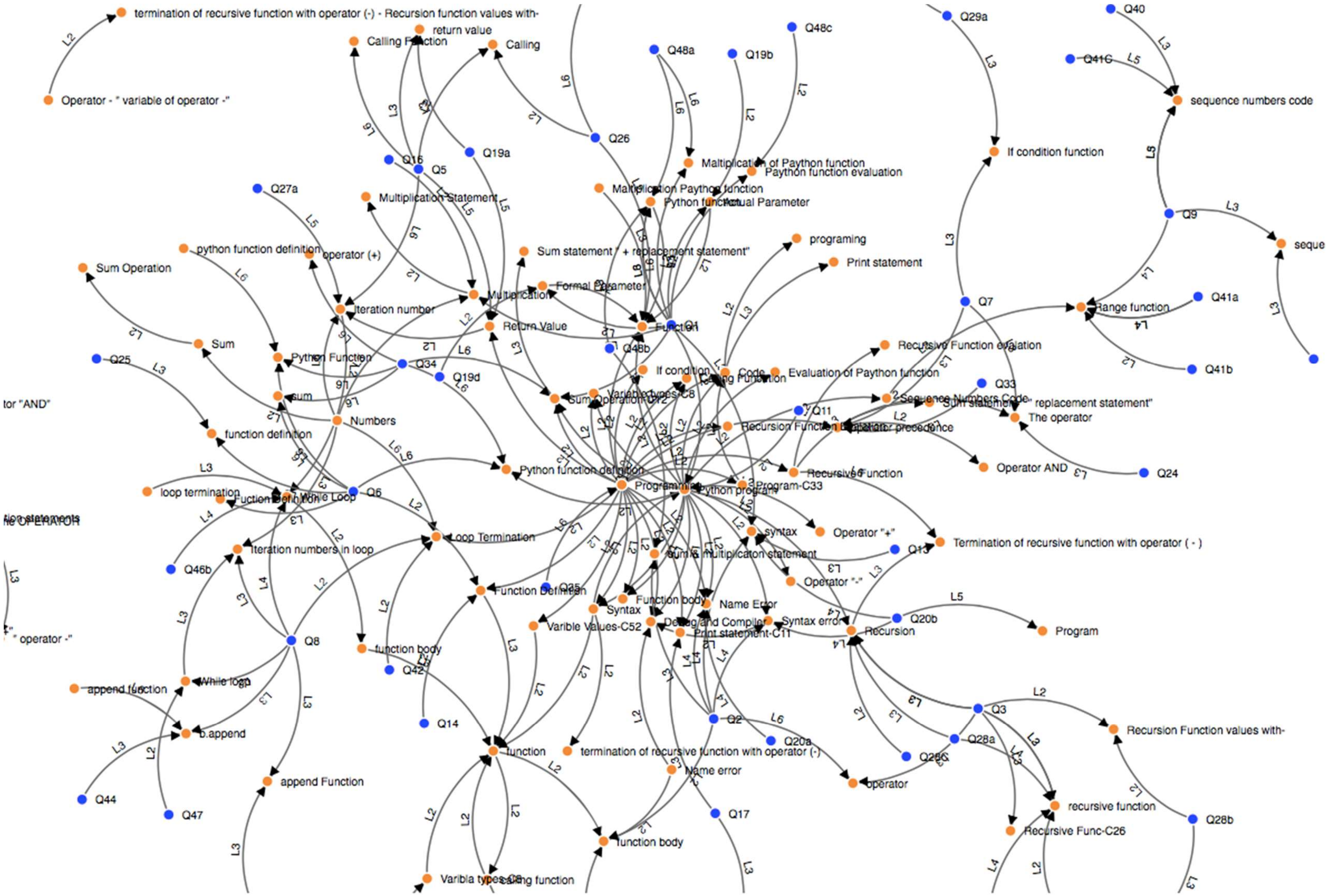
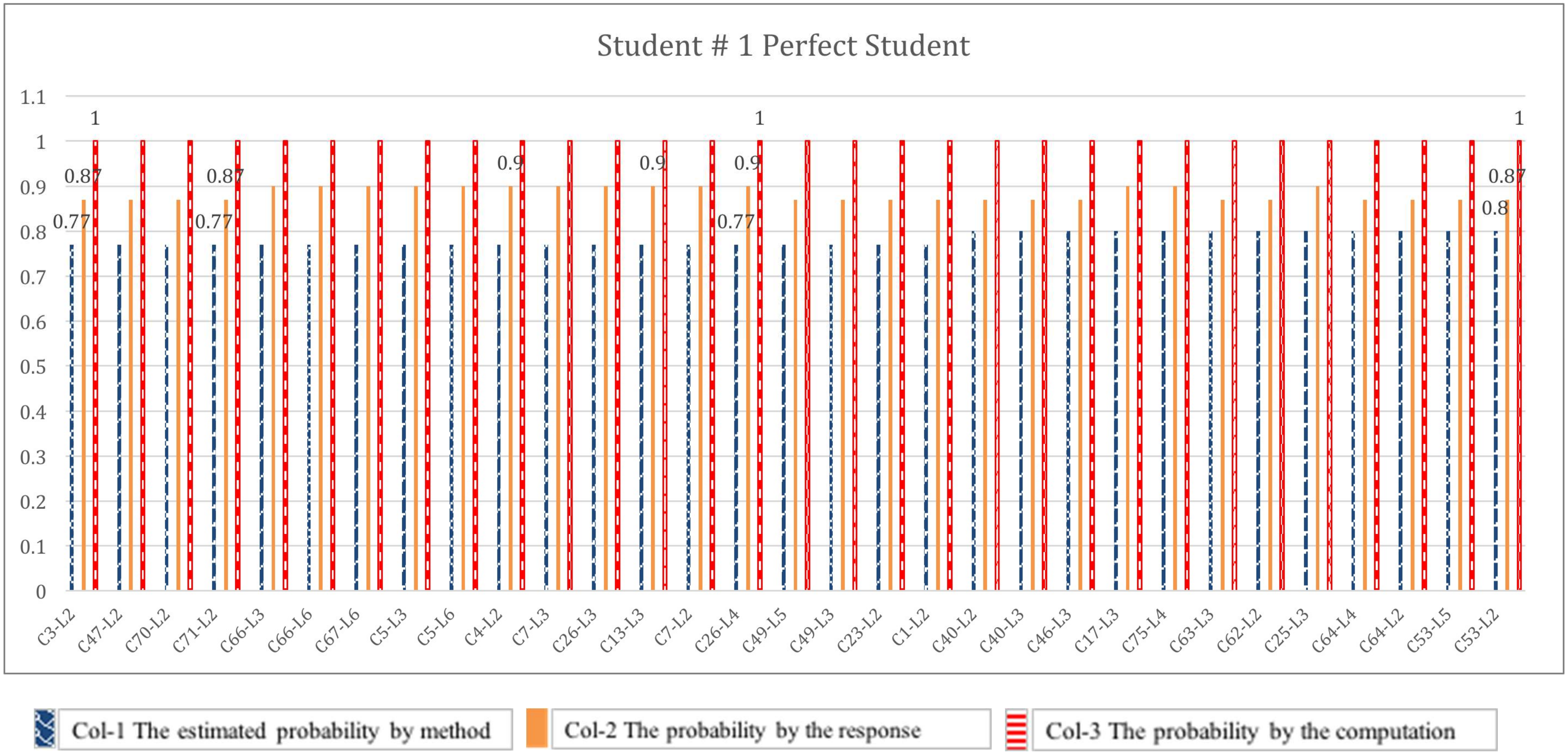


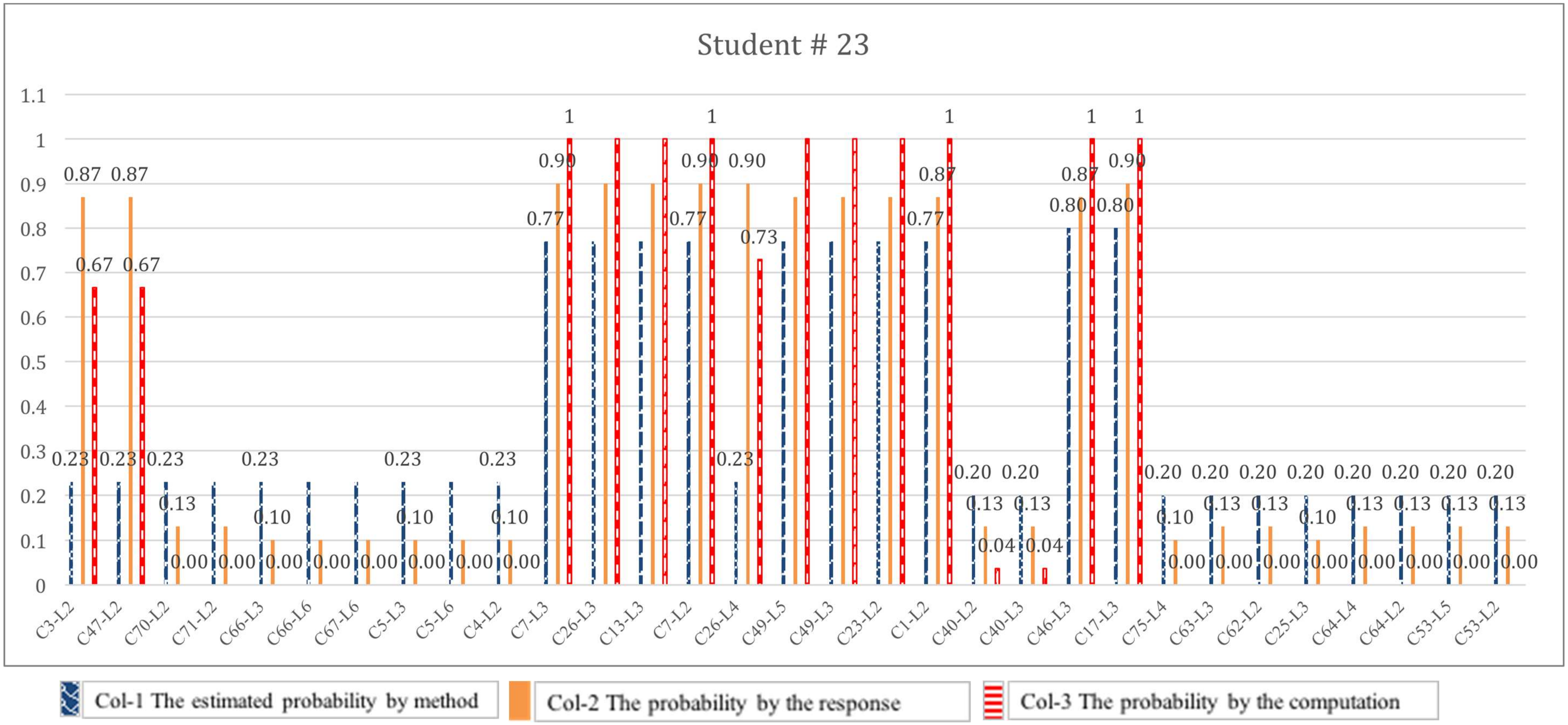
Figure 7.10 The Structure of the Concepts & the question Dependency of the Experiment 2



X: Concept # at the considered skill level

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

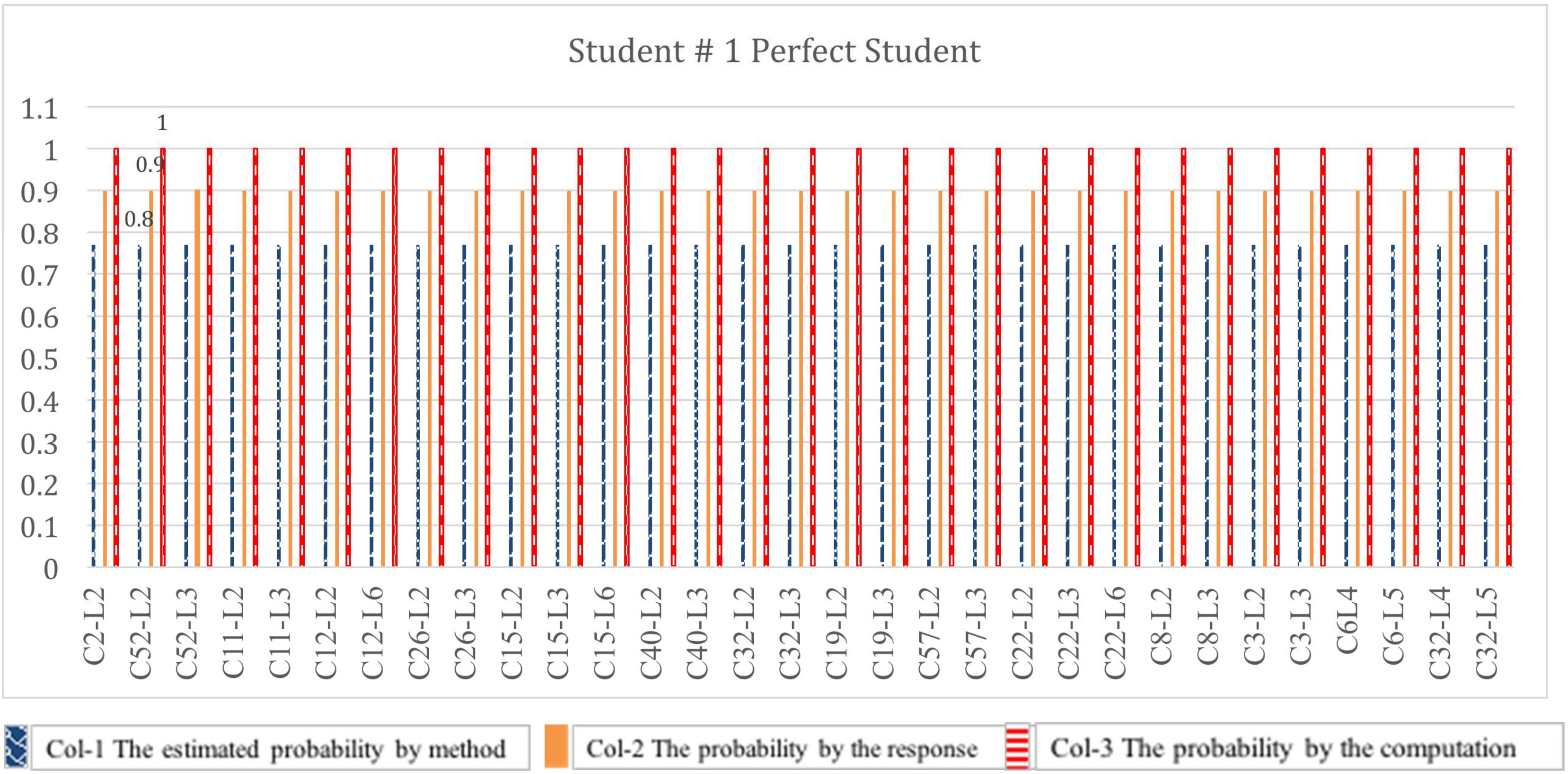
Figure 7.11 The Comparison Between the Probability of Knowing the Concepts Based on DS Method, Real Responses and Computed Probability by the Perfect Learner



X: Concept # at the considered skill level

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

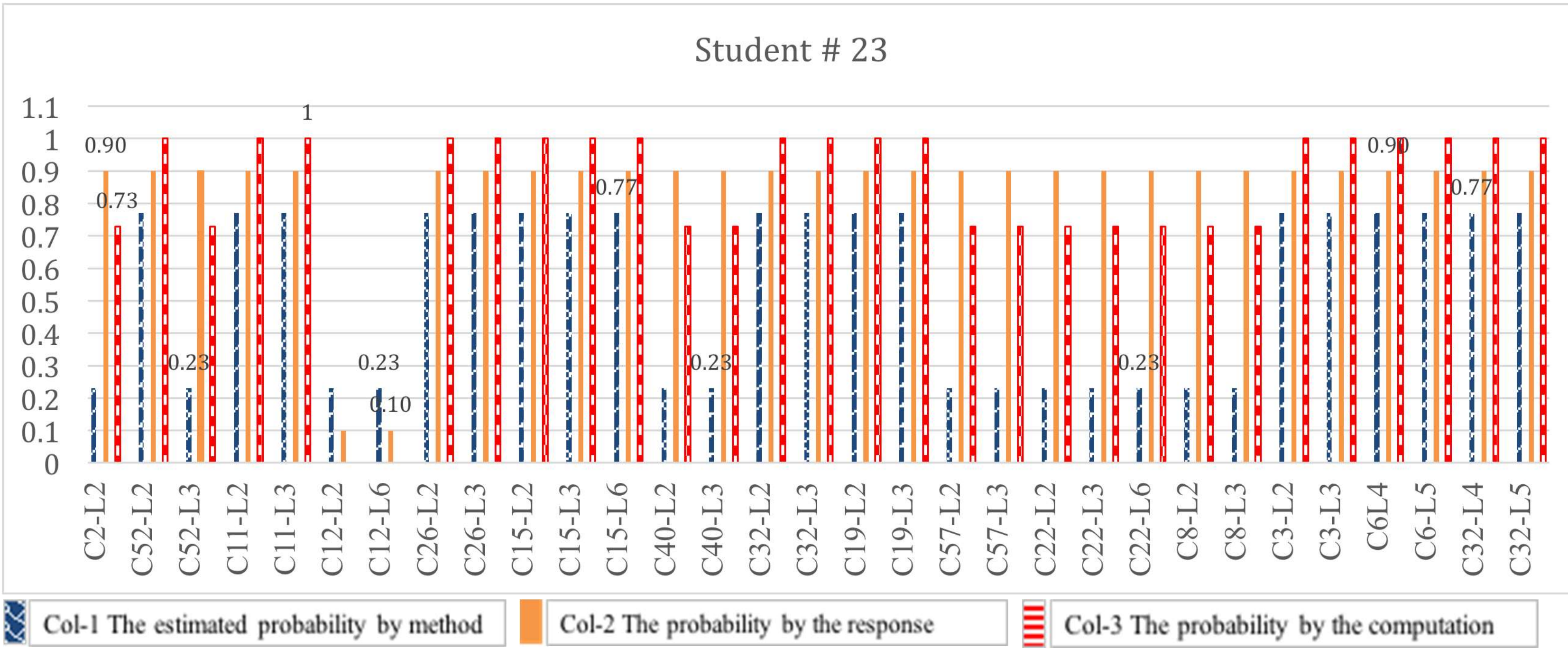
Figure 7.12 The Comparison Between the Probability of Knowing the Concepts Based on DS Method, Real Responses and Computed Probability of Randomly Selected Learner



X : Concept # at certain skill level

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

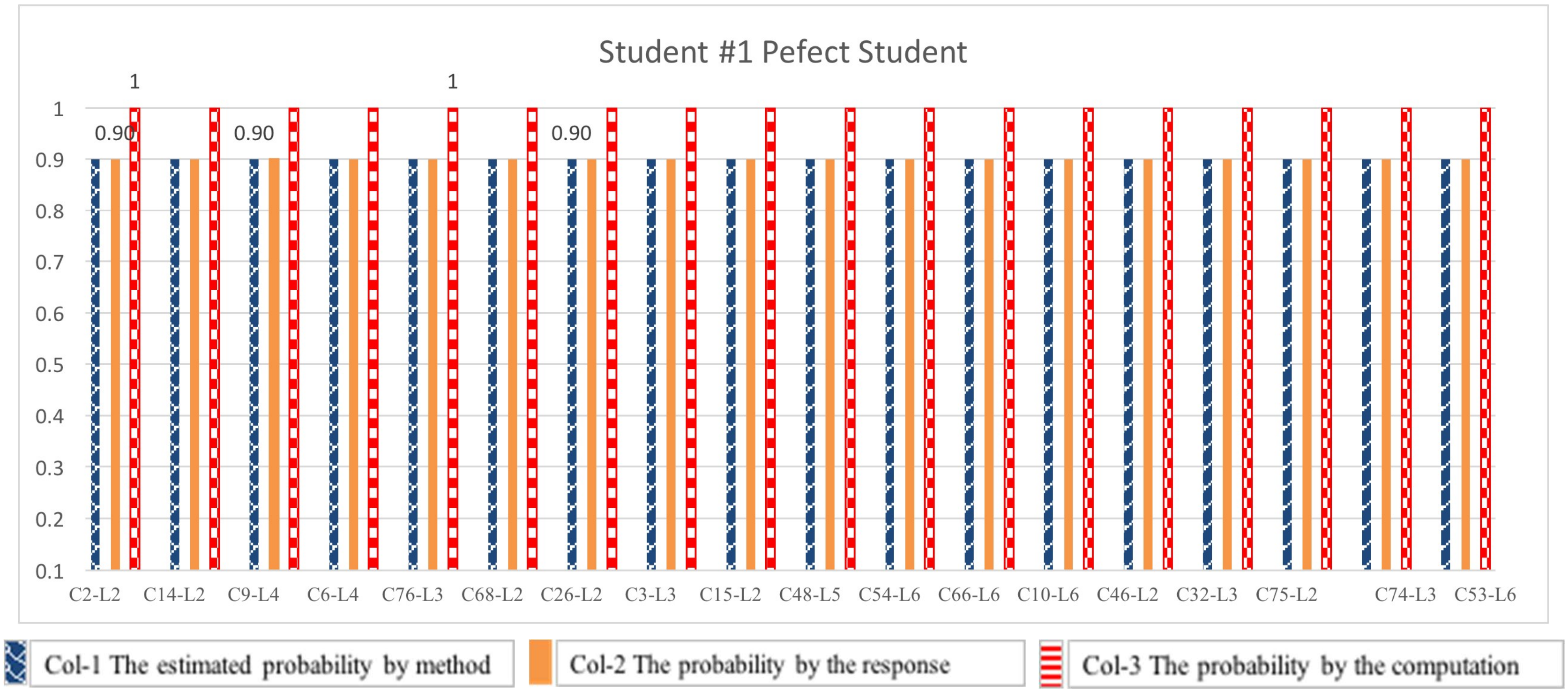
Figure 7.13 The Comparison Between the Probability of Knowing the Concepts Based on PS Method, Real Responses and Computed Probability of the Perfect Learner



X: Concept # at a certain skill level

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

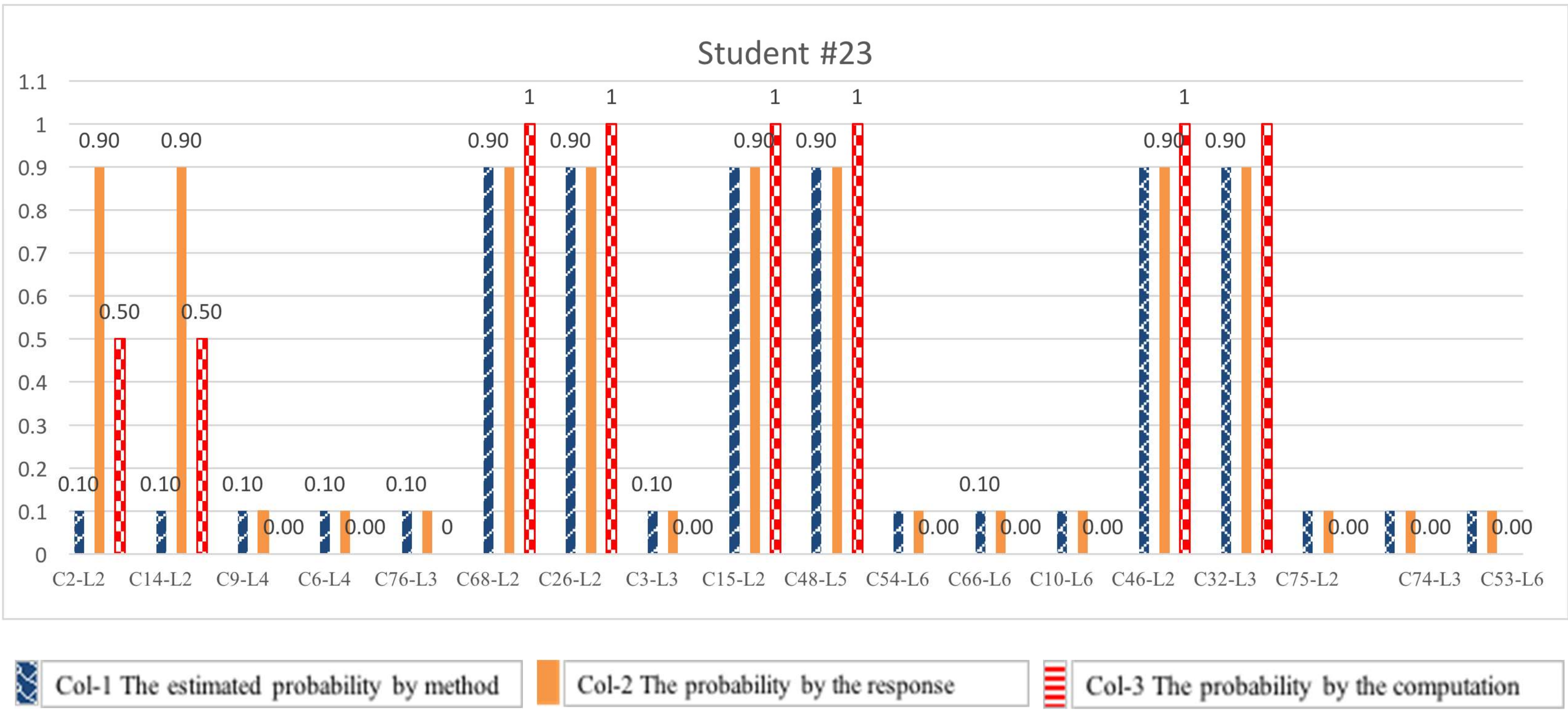
Figure 7.14 The Comparison Between the Probability of Knowing the Concepts Based on PS Method, Real Responses and Computed Probability of Randomly Selected Learner



X: Concept # at a certain skill level

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

Figure 7.15 The Comparison Between the Probability of Knowing the Concepts based on VS Method, Real Responses and Computed Probability of the Perfect Learner



X: Concept # at a certain skill level

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

Figure 7.16 The Comparison Between the Probability of Knowing the Concepts Based on VS Method, Real Responses and Computed Probability of Randomly Selected Learner

7.4.1 The Error Values in the Estimated Probability of Knowing the Concept Using the DS Method

1) If the concept at a certain skill level is tested by an indirect question, which is a regular open-ended question that has been prepared by the instructor and hasn't been concentrated in the skill level of the concept, then let the errors value be 0.2.

2) If the questions are multiple-choice and open question, then let the error value be 0.03+0.2 = 0.23. Thus,

3) If the response to question q_r is correct then the probability of knowing the concept C_j is $P(C_j|Q_r) = (1-g_r) = 1-0.23 = 0.77$ and the probability of not knowing the concept C_j is $P(C_j|Q_r) = g_r = 0.23$. "g_r" is a lucky guess error with the suggested value = 0.23.

4) If the response to a question q_r is incorrect, then the probability of knowing the concept C_j, which has been asked by the question q_r is $P(C_j|\overline{Q}_r) = C_r = 0.13$ and the probability of not knowing the concept C_j is $P(\overline{C}_j|\overline{Q}_r) = (1-C_r) = 0.77$, "$C_r$" is careless error with suggested value = 0.23.

7.4.2 The Probability Values of the Concept in DS by Direct Question

1) If the response to the question q asked about tested concept C_j is correct, then the probability of knowing the concept C_j is $P(C_j|Q_r) = (1-g_r) = 1\text{-}0.1=0.9$ and the probability of not knowing the concept C_j is $P(\overline{C}_j|Q_r) = g_r = 0.1$, g_r is a lucky guess error with suggested value = 0.1.

2) If the response to the question q_r is incorrect, then the probability of knowing the concept C_j, which has been asked by the question q_r is $P(C_j|\overline{Q}) = \mathcal{C}_r = 0.1$ and the probability of not knowing the concept C_j is $P(\overline{C}_j|\overline{Q}_r) = (1-\mathcal{C}_r) = 0.9$. "$\mathcal{C}_r$" is a careless error with suggested value = 0.1.

3) If the question is multiple-choice and it is a direct question, then let the probability of error be 0.03+0.1 = 0.13.

4) If the response is correct, then the probability of knowing the concept C_j is $P(C_j|Q_r) = (1-g_r) = 1-0.13 = 0.87$, and the probability of not knowing the concept C_j is $P(\overline{C}_j|Q_r) = g_r = 0.13$. " g_r " a lucky guess error with suggested value = 0.13.

5) If the response to a question q_r is incorrect, then the probability of knowing the concept C_j, which has been asked by the question q_r is $P(\overline{C}_j|\overline{Q}_r) = \mathcal{C}_r = 0.13$ and the probability of not knowing the concept C_j is $P(C_j|\overline{Q}) = (1-\mathcal{C}_r) = 0.87$. "$\mathcal{C}_r$" is careless error with suggested value = 0.13.

7.5 Some Studies of the Proposed TCS^2 Theory

7.5.1 Evaluation of the Test by Using a Parameter of Cognitive Skill Level

Educators have often tried to design an assessment sufficient to cover the broader objectives of a course. At the time, they have tried to assess the teaching material. This is not easy to measure. Moreover, the evaluation result is still limited to the methods of quantity percentage values of the topic, and this ignores the measurement of what skill of the concept has been learned. Does the quantity percentage value give a reliable measurement of achieving the course objective? Suppose we want to measure the outcome of the course, which has specific knowledge items that have to be learned aimed at a specific goal for the assessors to give the learners a letter grade at the end? Getting an A means the learner has met the criteria. What if the questions don't test the objective at all? How do we ensure that the test really reflects the original objective in the course? If the goal of the course is to teach learners how to apply algorithmic analysis, is it possible that the test itself doesn't test algorithmic analysis in Applying skill? At the end of the semester, will the test evaluation give the result that measures the analysis goal or design goal? There is no way to objectively measure that. In this measurement area, we add a new parameter, which is the Cognitive Skill Level, to the knowledge assessment. The Cognitive Skill Levels refer to levels, such as whether a learner has acquired the concept at the level of Understanding, or Applying, or Analyzing, or Evaluating, or Creating. Identifying the Cognitive Skill Level parameter of the concept in knowledge domain provides an accurate measurement value to the assessment. In this attempt, we provide knowledge assessment analysis methods that helps to design a proper test and that can measure exactly the covered knowledge of the course objective. However, measuring the concepts in the Concept Space, which includes the cognitive level parameter, allows the educators to design the test with optimal questions. The study of this benefit was published (Aboalela & Khan, 2016). We conclude that considering the cognitive level parameter helps the instructors avoid the overlap between questions that are introduced to the learners. Moreover, it helps to evaluate the efficiency of the assessment in three perspectives: the depth of the test, the overlap between the skills levels in the Concept

State, and the quality of the questions that measure most of the skills of the knowledge domain to achieve the course objective.

For proving this benefit, we used an actual test presented to 45 graduate learners in the course CS 46101 Design and Analysis of Algorithms. We call this test OQ. We used the same studied test in Experiment 1. We chose the test from a sorting algorithm chapter to evaluate it according to the proposed analysis methods. Therefore, the assessment domain is sorting algorithms, which we call Concept Space. The course objective is to Apply important algorithmic design paradigms and methods of analysis and the asymptotic performance of algorithms. The evaluated test involved 9 questions. Each question includes a specific tested concept at certain skill level. The depth of the test was evaluated by calculating the size of footprint of the skill level parameter. The analysis process was applied on OQ to find out the depth of the skill parameter in the tested concepts. Subsequently, we found out the overlap between the tested skills of the concepts. Finally, to find out the covered skills in the test we illustrate the skills plots of the concepts, counting, and percentage of tested and untested skill levels by OQ.

A. The Size of Footprint

The size of footprint is already discussed and calculated for the same test that has been studied in Experiment 1 in a previous section. We conclude that the prepared test by the instructor asked directly about 25 concepts. By adding the cognitive level as measurement parameter, we found that the test asked about 48 concepts at certain skills. Thus, considering a parameter of Cognitive Skill Level in the test management could help both to enhance the test's performance and to evaluate the assessments. The instructor could choose optimizing questions that asked about the minimum number of concepts but give more precise estimation of the learner knowledge.

B. The Overlapping Between the Skills in the Concepts States

The overlap between the skills means that two questions are asked for the same concept at the same skill level. The overlapping between the skill levels of the concept set in the Concept States is illustrated in Figure 7.17. There are 8 overlaps of the skills between the three Concept States corresponding to 83 skills of 53 concepts that exist in the Concept States. The percentage of overlaps is 13.56%. There are 4 overlaps between VS and DS, 1 overlap between DS and PS, and 3 overlaps between VS and PS, which means the test could be more efficient if it is implemented based on the proposed Concept State analysis. The instructor could implement the test from CLMCG, which presents the concepts in the Syllabus Domain, mapped together by prerequisite relation based on RBT skill levels.

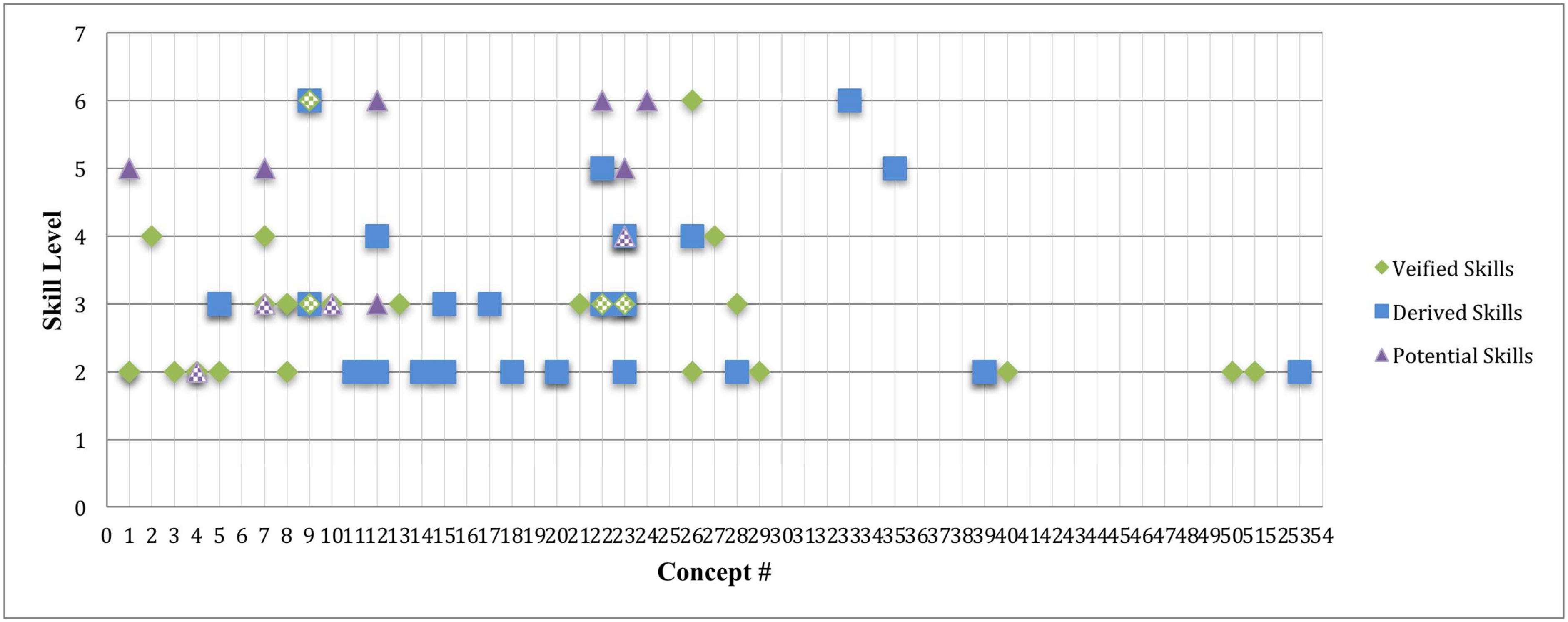


Figure 7.17 The Overlaps Between the Concepts in the Domain of Open Test

C. The Coverage of the Course Objective

We illustrate the covered skills in OQ by using a plot diagram. Figure 7.11 shows uncovered skills in the test. It should be known that each concept could be tested at many levels. Also, each question in OQ could test more than one skill and more than one concept. As observed in Figure 7.18, there are 24 uncovered skills of the Concept Space. Most of the uncovered skills are in the level 5, which is the evaluation level. Figures 7.19 and 7.20 show the counting and percentage of tested and untested skill levels by OQ respectively. The skills that have to be evaluated in the Concept Space are 83. In Figure 7.20, as evident, the highest tested skills are in levels 3 and 4 with 100%, which means the test covered all the concepts in the Concept Space of these two levels. The highest untested skills are in levels 5 with 70%. From figure 7.19 the number of untested skills of Level 5 is 14 of the 20 total skills that must be tested, which means there are still 6 more skills that should be measured. As known, the chapter objective is to achieve, Apply, Analyze, and create in the meaning of “write existing algorithms” but not to “evaluate” the algorithms. The unmeasured skills of the concepts are in the levels that don’t satisfy the course objective. Thus, the test given by the instructor, which we evaluated by using our analysis methods, is efficient to measure most skills of knowledge domain to achieve the course objective.

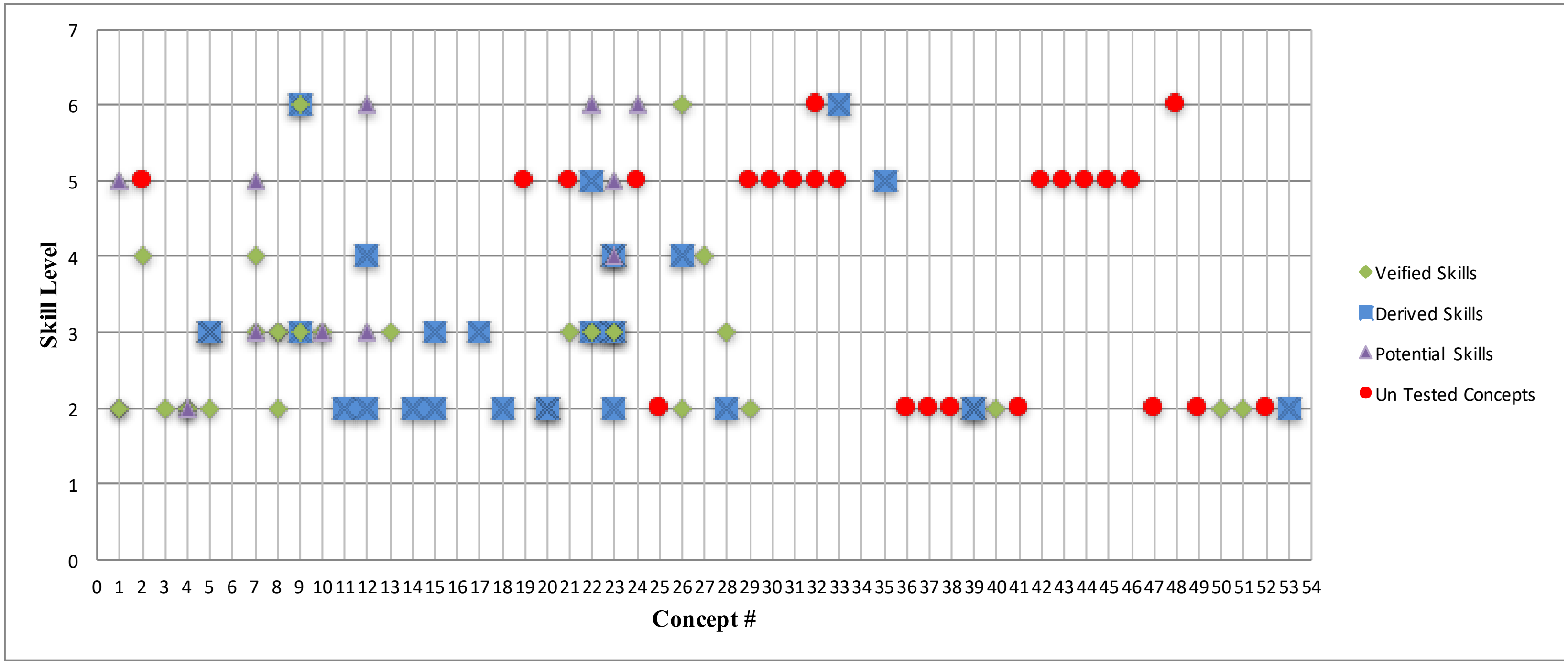


Figure 7.18 Uncovered Skills in the Concept States by Open Test

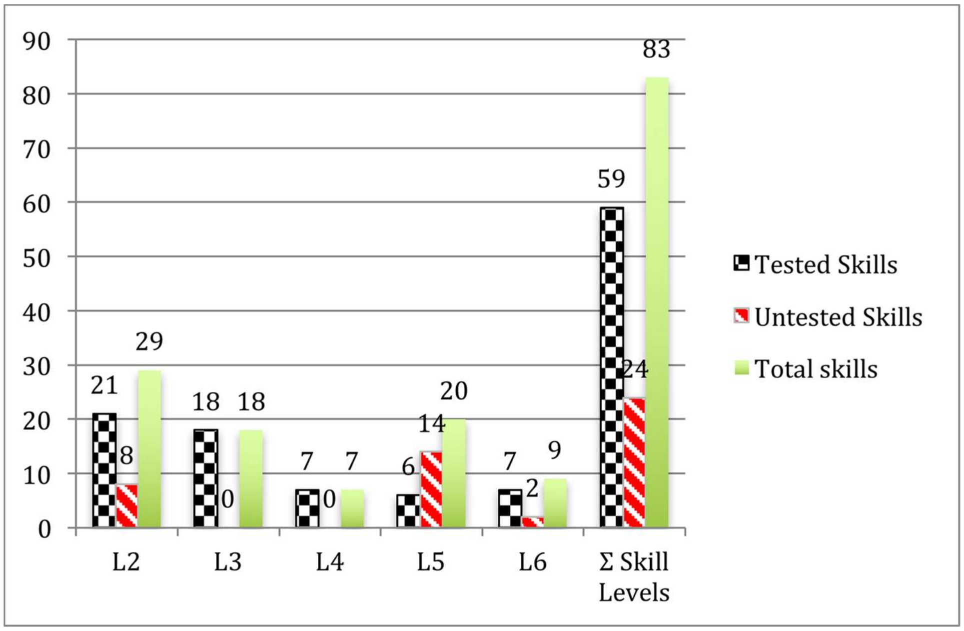


Figure 7.19 Counting of Covered and Uncovered Skills of the Concept Space

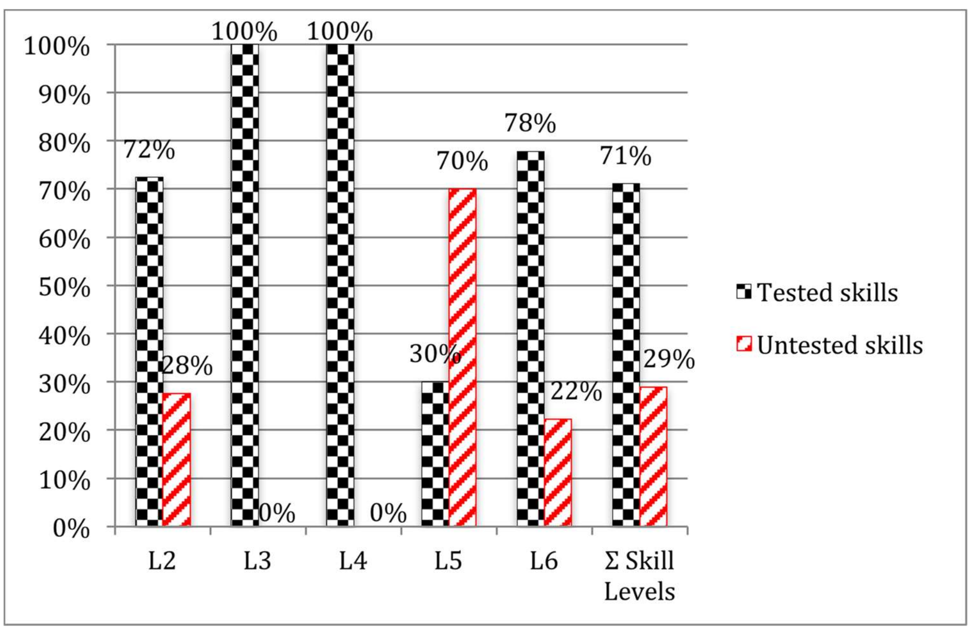


Figure 7.20 Percentage of Covered and Uncovered Skills of the Concept Space

7.5.2 Test Implementation

- We implement a sample of test connected with visualized Concept Space of study knowledge domain under study. This sample test was conducted to validate the proposed TCS^2 theory and to visualize the content of real book. In the test content, we visualize the test and the learner answers based on the ad concept according to the cognitive level mapped concept graph. The link is in[15]

D3JS for Visualization.

- D3JS is a JavaScript framework which lets us build data driven visualizations. D3JS focuses on binding data with DOM elements.
- I used D3JS tools to build a three-different interactive graph which deals with the concepts and level of understanding with learners, it also shows the concepts and relation with other concepts as well.
- The three different graphs, are
    - Bloom graph
    - Ontology graph
    - Hierarchy graph
- Bloom graph:
    - The graph is an interactive build with d3js force graph layout. I built a legend for the force graph.

---

[15] http://rania.medianet.cs.kent.edu:8080/Project

- The graph can reach to two levels. The initial level, as mentioned, is a force graph layout with nodes and links. The Nodes are circles, which represent concepts. The Links are curved lines, which represent value between concepts.
- Technically: - I used the d3.csv () function to read the data.csv file which consists of three columns of data (Two columns are concepts and the other is value).
- Once the program reads the file, it creates an array links with all the data, and Loops through links by using force function to append nodes and links with respective value.
- Second level: - It consisted of the selected node as center and all the relatives of the node around it with their levels.
- On double click in the first level, we removed the current SVG parent element and started creating the other SVG element in the same place with neighborhood value, which was fetched from the neighbor function neighbors().

- Ontology graph: - The Ontology graph follows the same force graph methodology as Bloom's graph with changes in nodes from to square circle shape. The links are representing different values and different models respectively.
- Hierarchy graph: - This is a tree structured d3js graph which explains data structures. This graph also consists of nodes and links, but, unlike a force graph, it follows hierarchy structure, which helps in represent the parent-child relationship.

JSP

- JSP (Java Server Pages) is a language which binds Java with HTML and makes GUI. We created the Index.jsp which starts the Home Page for learner examination.
- In the Home page, the project allowed users to register, which added a record in database with details of the Mail Id, username and password. The role could be selected with drop down list.
- The project followed for interaction between JSP and Database MySQL. Steps involved in the project are:
    - It verified user credentials with the user details table from database with three options. Admin, Grader and Learner pages.
    - The Admin page allowed for user to interact with question table, where we could add questions to the exam paper.
    - The Grader page allowed for correction of the examination, and for the ability to see the graphs.
    - The Learner page allowed users to attend the examination and submit it.

MySQL

It was the Database to store records with the interactive tool (phpMyAdmin).

IDE

- The Sublime Text for d3js visualization.
- The Eclipse for Examination project.
- The Apache Tomcat for Web Services.

CHAPTER 8

Conclusion and Future Work

In this research study, an Assessment Theory of Cognitive Skills in Concept Space (TCS^2) was proposed. The new analysis method used to map the concepts in the knowledge domain in one space is called CLMCG, and the study includes instances of its application. The research study also includes a new perspective on the result of the assessments revealing “Concept States”. The study indicates that the learners could have the result as six sets of the Concept States. These proposed Concept States are tested and validated. The significant value of the Concept States is that they reveal the exact concepts within the learner knowledge, as well as the exact skill level of the concept; both of which provide the additional benefit of increased precision of the assessment feedback. A sampling of applications that provide evidence of the efficacy of the introduced methods is also provided.

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.

The results of this study add to the precision of knowledge assessments by adding the parameter of Cognitive Skill Levels to the assessment protocols, and by increasing the amount of estimated knowledge, despite being able to prepare the tests using the minimum number of specified concepts, pertaining to the maximum number of concepts in the domain.

The study provides computational formulae to ascertain the precise probability of knowing each concept in the assessment domain of the assessed individual. Moreover, the combination of the probability of all the concepts are used to estimate the probability of the learning objectives involving these concepts, as well as the learning objectives in the domain.

The fundamentals of knowledge assessment are demonstrated by an example of learners' knowledge. In future work, these fundamentals could be used for further knowledge assessment, such as the assessment of understanding and comprehension in the human brain. The assessments could also be used to track the brain development and could include factors such as the background of the individual, his/her work in the present, and the other relationships affecting learning such as social media and social networks. This research study used basic and simple methods that may be used for assessment in any field, and is not limited to assessing learner learning. The computation of the probability of the concepts could be used in the fields of healthcare in its reliance on simple and clear mapping methods.

References

Aboalela, R., & Khan, J. (2015). Visualizing concept space of course content. *IEEE 7th International Conference on Engineering Education (ICEED)* (pp. 160-165). Japan: IEEE .

Aboalela, R., & Khan, J. (2016). Are we asking the right questions to grade our students in a knowledge-state space analysis? *2016 IEEE 8th International Conference on Technology for Education (T4E 2016).* Mumbai.

Albert, D., & Held, T. (1999). Component based knowledge spaces in problem solving and inductive reasoning. 15-40.

Anderson, L. W., Krathwohl, D., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., . . . Wittrock, M. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* New York, U.S.A: Longman.

Auzende, O., Giroire, H., & Le Calvez, F. (2009). Using Competencies to Search for Suitable Exercises. *International Conference on Advanced Learning Technologies (ICALT)* (pp. 661-665). Latvia: IEEE.

Chen, Y. (2013). Adaptive Non-graded Assessment Based on Knowledge Space Theory. *Advanced Learning Technologies (ICALT), EEE 13th International Conference on.* Beijing.

Doignon, J. P., & Falmagne, J. C. (1999). *Knowledge Spaces.* Berlin: Springer.

Doignon, J.-P. (1994). *Knowledge spaces and skill assignments.* (G. Fischer, & D. Laming, Eds.) New York: Springer-Verlag.

Dowling, C. E., Hockemeyer, C., & Ludwig, A. H. (1996). Assessment and Training Using the Neighbourhood of Knowledge States. In I. T. Systems (Ed.), *In Claude Frasson□ Gilles Gauthier□ and Alan Lesgold. 1086*, pp. 578-586. Berlin: Springer Verlag.

Falmagne, J. C., Cosyn, E., Doble, C. T., & Uzun, H. (2007). Assessing mathematical knowledge in a learning space: Validity and/or reliability. *In Annual Meeting of the American Educational Research Association (AERA).*

Falmagne, J., Cosyn, E., Doignon, J., & Thiery, N. (2003). The assessment of knowledge, in theory and in practice. *International Conference on Integration of Knowledge Intensive Multi-Agent Systems* (pp. 609 - 615). IEEE.

Falmagne, J.-C., & Doignon, J.-P. (1988). A Markovian procedure for assessing the state of a system. *Journal of Mathematical Psychology, 32*(3), 232–258.

Ford, W., & Topp, W. (2001). *Data Structures with C++ using the STL* (2nd Edition ed.). Pearson.

Heller, J., Mayer, B., Hockemeyer, C., & Albert, D. (2005, March). Competence-based knowledge structures for personalised learning: distributed resources and virtual experiments. *International Journal on E–Learning*, 14-16.

Heller, J., Steiner, C., Hockemeyer, C., & Albert, D. (2006). Competence-based knowledge structures for personalised learning. *International Journal on E-learning, 5*, 75–88.

Hockemeyer, C. (1993). *Knowledge diagnosis on knowledge spaces. Thesis, Braunschweig, Germany.*

Khan, J., & Hardas, M. (2013). Does sequence of presentation matter in reading comprehension? A Model based Analysis of Semantic Concept Network Growth during Reading, Semantic Computing (ICSC. *IEEE Seventh International Conference* (pp. 444-452). IEEE.

Khan, J., Hardas, M., & Ma, Y. (2005). A Study of Problem Difficulty Evaluation for Semantic Network Ontology Based Intelligent Courseware Sharing. *IEEE/WIC/ACM International Conference on Web Intelligence* (pp. 426-429). WEB Intellegence.

Khan, J., Ma, Y., & Hardas, M. (2006). Course composition based on semantic topical dependency. *International Conference on web Intelligence IEEE/WIC/ACM* . IEEE.

Lee, P. (2012). *Bayesian Statistics: An Introduction* (4th Edition ed.). Willey.

Marte, B., Steiner, C. M., Heller, J., & Albert, D. 4. (2008). Activity-and taxonomy-based knowledge representation framework. *International Journal of Knowledge and Learning, 4*(2/3), 189-202.

McGrew, T. J. (2005). Retrieved from http://homepages.wmich.edu/~mcgrew/Bayes8.pdf

Nafa, F., & Khan, J. (2015). Conceptualize the Domain Knowledge Space in the Light of Cognitive Skills. *In Proceedings of the 7th International Conference on Computer Supported Education. 1*, pp. 285-295. SCITEPRESS-Science and Technology Publications.

Nafa, F., Khan, J. Othman, & Babour, A. (2016 a). Discovering Bloom Taxonomic Relationships between Knowledge Units Using Semantic Graph Triangularity

Mining," . *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)* (pp. 224-233). Chengdu: IEEE .

Nafa, F., Khan, J., Othman, S., & Babour, A. (2016 b). Mining Cognitive Skills Levels of Knowledge Units in Text Using Graph Tringluarity Mining. *Web Intelligence Workshops (WIW), International Conference (IEEE/WIC/ACM)* (pp. 1-4). IEEE.

Novak, J. a. (2008). *The Theory Underlying Concept Maps and How to Construct and Use Them.* Florida Institute for Human and Machine Cognition, Technical Report IHMC CmapTools 2006-01 Rev 2008-01, Florida.

Reimann, P., Kickmeier-Rust, M., & Albert, D. (2013, May). Problem solving learning environments and assessment: A knowledge space theory approach. *Computers & Education, 64*, 183–193.

Rohatgi, V., & Saleh, E. (2015). *An Introduction to Probability and Statistics* (3ed ed.). Wiley.

Stahl, C. (2011). *Knowledge Space Theory. 2008-06-18).[2009-12-22].* Retrieved from http://www2.uaem.mx/r-mirror/web/packages/kst/vignettes/kst.pdf

Zwillinger, D., & Kokoska, S. (2000). *Standard probability and Statistics tables and formulae.* Chapman & Hall/CRC.