S V E U Č I L I Š T E U S P L I T U
FAKULTET ELEKTROTEHNIKE, STROJARSTVA I BRODOGRADNJE

POSLIJEDIPLOMSKI DOKTORSKI STUDIJ ELEKTROTEHNIKE I INFORMACIJSKIH TEHNOLOGIJA

KVALIFIKACIJSKI ISPIT

STUDENT MODELLING USING BAYESIAN KNOWLEDGE TRACING

Ines Šarić-Grgić

Split, ožujak 2023.

UNIVERSITY OF SPLIT
FACULTY OF ELECTRICAL ENGINEERING,
MECHANICAL ENGINEERING AND NAVAL ARCHITECTURE

POSTGRADUATE STUDY OF ELECTRICAL ENGINEERING
AND INFORMATION TECHNOLOGY

QUALIFICATION EXAM

STUDENT MODELLING USING BAYESIAN KNOWLEDGE TRACING

Ines Šarić-Grgić

Split, March 2023

CONTENTS

1. INTRODUCTION 1
2. STUDENT MODELLING 3
3. VANILLA BAYESIAN KNOWLEDGE TRACING (BKT) 6
4. METHODOLOGY 8
5. BKT ENHANCEMENTS.. 11
5.1. The architectural and educational context-based enhancements (RQ1) 11
5.2. Computational methods (RQ1) 13
5.3. Evaluation approaches (RQ2) 14
6. CONCLUSION 17
BIBLIOGRAPHY... 19
ABBREVIATION LIST 26
ABSTRACT... 27
SAŽETAK 28

1. INTRODUCTION

In the traditional face-to-face learning environment, a teacher makes conclusions about learning progress based on students' responses to different questions and assignments in the classroom. During this process characterized by uncertainty, the teacher's task is to estimate the student's current knowledge level and adapt the teaching approach. Since students differ in prior knowledge and learning abilities, their knowledge assessment can be challenging even for experienced teachers. A more appealing environment is the tutoring environment as one-on-one tutor-student interaction in which technology mimics human teachers. Such environment enables personalized and adaptive tutoring, taking into account potential difficulties and misconceptions.

The researchers of the interdisciplinary field of cognitive science, artificial intelligence, and educational technology have computerized teaching and learning since the 1960s by developing various types of educational platforms, e.g. Computer Assisted Instruction (CAI) [1], Intelligent Tutoring Systems (ITS) [2], Intelligent Learning Environments (ILE) [3], Adaptive Instructional Systems (AIS) [4], etc. While all these platforms target adaptive and intelligent behaviour, the extensive research of ITS over the years gathered the scientific community and represents a valuable research background. The ITS community agreed on the standard architecture and set the student module (includes information about what to learn) as one of the four basic structural components of the generalized ITS [5]–[9]. Other modules include the domain knowledge module (includes information about what to learn), the tutoring module (how to teach), and the communication module (tutor-student interface). The main quality of an ITS is its ability to observe and interpret student behaviour to infer the preferences and needs of an individual student. The student model defined in the student module is an essential component of the ITS that provides adaptive and personalized tutoring. It enables a comprehensive representation of student knowledge and affects the quality of the other ITS components.

Besides the previous, other widely used educational platforms are Learning Management Systems (LMS) such as Moodle and Massive Open Online Courses (MOOC). Although these platforms do not incorporate adaptive and intelligent behaviour, their popularity offers valuable testing ground to improve the teaching, learning and testing processes. They do track various student interactions, so they are often the focus of research in the field of Educational Data Mining (EDM) and Learning Analytics (LA).

The main challenge of the student model is the knowledge inference process that aims to estimate the knowledge level during the learning process. Bridging the gap between the student's input to the system and the system's conception and representation of a student's knowledge is known as *knowledge tracing* [10]. Over the years, researchers used different Machine Learning (ML) techniques to model student knowledge. According to previous review studies on student modelling [11]–[23], Bayesian networks are continuously investigating ML technique and the *Bayesian Knowledge Tracing (BKT)* [10] based on the Hidden Markov Model is the most representative and still state-of-the-art approach. The BKT advantages include its *simplicity* in definition, *ability* to infer student knowledge, parameter *interpretability* and *applicability* to datasets limited in size. Since the vanilla term refers to technology not customized or updated from its standard form, the BKT model first proposed by Corbett and Anderson [10] is further called the *vanilla BKT model.*

Based on the vanilla BKT model, various enhancements have been investigated in the literature, however, there is no systematic and up-to-date review on the family of BKT models. In the following sections, we present the review study of the BKT enhancements based on uniquely proposed criteria that apply the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines [24]. The review study aims to answer the following Research Questions (RQ):

RQ1: What has been proposed in the literature to enhance the vanilla BKT model since its emergence in 1995?

RQ2: Which evaluation approaches, including educational platforms and performance measures were part of the research on the BKT enhancements?

The remainder of the qualification exam discusses previous student modelling studies, presents brief description of the vanilla BKT model, the methodology used to review the BKT enhancements, the results and finally, the conclusion.

2. STUDENT MODELLING

The previous research generally refers to *overviews* [16], [20]–[23] and *reviews* [11], [13]–[15], [17], [18]**.** Only a few studies used *systematic methodology* to investigate the literature [12], [19]. Table 2.1 provides the complete list of proposed taxonomies of student modelling approaches in descending chronological order. It consists of (i) the research study reference, (ii) the specificity of the study, and (iii) proposed student modelling taxonomy. The BKT-related student modelling approaches are emphasized in italic in Table 2.1.

Table 2.1: Proposed taxonomies of student modelling approaches.

Research study	**Specificity of the study**	**Proposed taxonomy of student modelling approaches**
Liu et al (2021)	The review from the technical point of view	*Probabilistic models;* Logistic models; Deep learning-based models
Ramirez Luelmo et al (2021)	The systematic review of machine learning techniques (2015-2020)	*Bayesian Knowledge Tracing;* Deep Knowledge Tracing; Long-Short Term Neural Networks; Bayesian Networks; Support Vector Machines; Dynamic Key-Value Memory Network; Performance Vector Analysis
Pelanek (2017)	The review focused on the macro adaptive behaviour (curriculum sequencing) (2014-2020)	*Bayesian Knowledge Tracing;* Logistic models
Anouar Tadlaoui et al (2016)	The review focused on the Adaptive Educational Systems	Overlay; Stereotype; Machine Learning*;* Plan Recognition; Differential; Perturbation; *Bayesian Networks*
Sani et al (2016)	The review focused on the Intelligent Tutoring Systems (2010-2015)	*Bayesian Knowledge Tracing;* Fuzzy Logic; Overlay; Differential; Perturbation; Constraint-based; Machine Learning; Stereotype
Kurup Et Al (2016)	The overview focused on the Intelligent Tutoring Systems	Overlay; Bayesian Network; Correct First Attempt Rate; Performance Factor Analysis; Tabling; *Bayesian Knowledge Tracing*
Zafar & Ahmad (2013)	The review of the student modelling approaches under uncertain conditions	Student modelling using statistical reasoning (Bayesian Networks, Reasoning using Certainty Factors); Fuzzy Modelling
Pavlik et al (2013)	The review focused on the Intelligent Tutoring Systems	Overlay models (Rule Space models, *Model Tracing models*, Constraint-based models); Knowledge Space models; Dialogue models; Programmed Branching; State and Trait
Chrysafiadi & Virvou (2013)	The systematic review focused on the Adaptive Educational Systems (2002-2012)	Overlay; Stereotypes; Perturbation; Machine Learning*;* Cognitive Theories; Constraint-based Model; Fuzzy Modelling; *Bayesian Networks*; Ontology-based Modelling
Harrison & Roberts (2012)	The overview of student modelling techniques for application in serious games	*Knowledge Tracing;* Performance Factor Analysis; Matrix Factorization
Desmarais & Baker (2012)	The overview of the most successful and widely used approaches focused on the macro adaptive behaviour (curriculum sequencing)	Tutors for problem-solving and solution analysis (Cognitive tutors and Constraint-based modelling); Content sequencing tutors (Models of skills - Bayesian Networks and graphical models, IRT and Latent Trait models, Latent cousins DINA, NIDA, DINO, NIDO, *Bayesian Knowledge Tracing*, Models without hidden nodes)
Vandewaetere et al (2011)	The overview of the parameters that are included in the student model when developing adaptive learning environments	Stereotypes; Feature-based modelling; Combination of stereotypes and feature-based modelling; Other approaches (Constraint-based modelling and Modelling of misconceptions)
Brusilovsky & Milan (2007)	The overview focused on the Adaptive Hypermedia and Adaptive Educational Systems	Overlay; Uncertainty-based modelling

Each study in Table 2.1 gave overview of student modelling approaches related to specific educational platforms, adaptive behaviour or used techniques. While the first review recognized the overlay and uncertainty-based student modelling approaches [23], later studies proposed more extensive taxonomies. In the early review studies, ML techniques were already recognized as the basis for different student modelling approaches [20], [21]. Over the time, the use of new techniques complemented the previous student modelling taxonomies [11], [12]. However, the new taxonomies of student modelling approaches are still adopted, and there is *no consensus* on the correct taxonomy.

As the subfield of artificial intelligence, ML works on algorithms that enable machines to learn through experience and using data [25]. ML techniques used for student modelling offered new ways to enhance the adaptiveness and intelligence of educational platforms. Those identified in the already mentioned research studies include the Bayesian Networks, the Logistic Regression, the Neural Networks, the Support Vector Machines, the Fuzzy Logic, and the Matrix Factorization (Table 2.2).

Table 2.2: ML techniques identified in the research on student modelling.

Research study	**Bayesian Networks**	**Logistic Regression**	**Neural Networks**	**Support Vector Machines**	**Fuzzy Logic**	**Matrix Factorization**
Liu et al (2021)	+	+	+	-	-	-
Ramirez Luelmo et al (2021)	+	+	+	+	-	-
Pelanek (2017)	+	+	-	-	-	-
Anouar Tadlaoui et al (2016)	+	-	-		-	-
Sani et al (2016)	+	-	-	-	+	-
Kurup Et Al (2016)	+	+	-	-	-	-
Zafar & Ahmad (2013)	+	-	-	-	+	-
Pavlik et al (2013)	+	-	-	-	-	-
Chrysafiadi & Virvou (2013)	+	-	-	-	+	-
Harrison & Roberts (2012)	+	+	-	-	-	+
Desmarais & Baker (2012)	+	-	-	-	-	-
Vandewaetere et al (2011)	-	-	-	-	-	-
Brusilovsky & Milan (2007)	+	-	-	-	+	-

Moreover, Liu et al. [11] summarized the student modelling approaches as probabilistic, logistic, and deep learning-based models. The *probabilistic* models, such as BKT, are based on Bayesian Networks and assume that the learning process follows a Markov process, which uses the observed states to estimate the student's hidden knowledge states. The *logistic regression* models, such as Learning Factor Analysis [26] and Performance Factor Analysis [27], involve the mathematical function of learning parameters and logistic regression to predict the probability of mastery. The last group of knowledge tracing approaches, the *deep learning* models, are based on neural networks, and they have been introduced in recent years [28].

Ramirez Luelmo et al. [12] investigated ML techniques employed in student modelling from 2015 to 2020. Their research results indicate the most common ML techniques as BKT (18 research studies), Deep Knowledge Tracing (13 r.s.), Long-Short Term Neural Networks (12 r.s.), Bayesian Networks (11 r.s.), Support Vector Machines (7 r.s.), Dynamic Key-Value Memory Networks (7 r.s.), and Performance Factor Analysis (6 r.s.).

Overall, Bayesian networks are the continuously investigated ML technique used for student modelling, and vanilla BKT based on the Hidden Markov Model is the most representative. Since the vanilla model is one of the first and still state-of-the-art approaches, researchers in the field have often recognized it as a unique student modelling approach.

As for the model preference, there is no general agreement on the choice between the probabilistic and the logistic models. The researchers often prefer one model but provide no rationale behind their choices [13]. On the other side, the apparent accuracy improvement of deep learning-based models over BKT was due to the high dimensional hidden space and ability to observe interleaved skills in a single model [29]. The comparison between Neural Network-based research and the vanilla BKT model revealed that simply enabling the forgetting parameter of the vanilla model led to a performance close to deep knowledge tracing on several datasets [30], [31]. The intuitive knowledge inference process of knowledge tracing makes the BKT model different from the Neural Network-based model, which does not have insight into it and hence lacks interpretability.

This work differs from other literature reviews on several accounts since it focuses on the probabilistic BKT models, systematically covers the research works published since the introduction of BKT in 1995 up to the most recent research and reviews the BKT enhancements and evaluation approaches.

3. VANILLA BAYESIAN KNOWLEDGE TRACING (BKT)

The vanilla BKT model is one of the first ML knowledge tracing models introduced by Corbett and Anderson [10]. It is considered as the first significant milestone of the EDM research field [32].

The vanilla model results from the work on the ACT Programming Tutor and reflects the ACT-R cognitive theory (Adaptive Control of Thought – Rational) [33], which states that mastering a complex skill implies its components. Also, it reflects the use of the Bayesian computational procedure identified in Atkinson's work [34].

A Bayesian Network is a probabilistic graphical model for representing knowledge about an *uncertain domain* where each node corresponds to a random variable and each edge represents the conditional probability for the corresponding random variables. A Hidden Markov Model is a special Bayesian Network used to trace not directly observable events by using a sequence of directly observable ones. In BKT, *student knowledge* is represented as a *hidden* node, while *student performance* is represented as an *observable* node. Both types of nodes are assumed to be binary, including the unlearned *(u)* and learned knowledge states *(l)* and the correct *(c)* and incorrect performance states *(i)*.

Figure 3.1 shows the hidden student knowledge nodes kc_t, $t \in \{1, 2,...T\}$ and observable student performance nodes sp_t, $t \in \{1, 2,...T\}$ of the vanilla model HMM's as circles and rectangles.

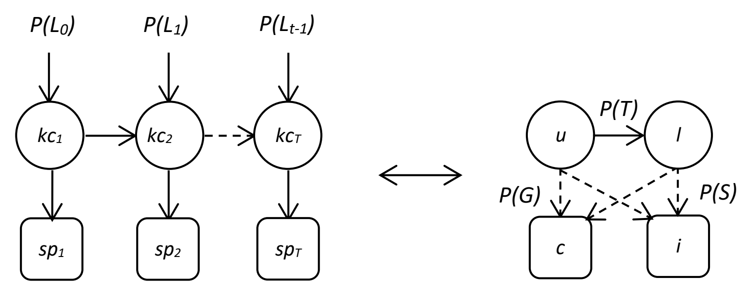

Figure 3.1: The vanilla BKT model and its instantiation process.

While $P(L_0)$ is the initial probability of knowledge before any opportunity of applying it (prior knowledge), the arrows between the circles indicate the *transition probabilities,* and the dashed ones (from the circles to the rectangles) indicate the *emission probabilities*. The transition probabilities refer to the probability $P(T)$ of a knowledge transitioning from unlearned state *u* to learned state *l* and to the probability $P(F)$ of forgetting a previously known knowledge which is assumed to equal zero in the vanilla model. The model defines emission probabilities by guessing the probability of correctly answering unlearned knowledge $P(G)$ and the slip probability of making a mistake when answering a learned knowledge $P(S)$.

Figure 3.2 shows the complete set of vanilla model parameters consisting of $P(L_0)$, $P(T)$, $P(G)$, and $P(S)$ in a matrix form.

Priors	
learned	$P(L_0)$
unlearned	1- $P(L_0)$

Transitions		
	to learned	to unlearned
from learned	1	0
from unlearned	$P(T)$	1- $P(T)$

Observations		
	correct	incorrect
learned	1- $P(S)$	$P(S)$
unlearned	$P(G)$	1- $P(G)$

Figure 3.2: BKT parameters in a matrix form [35].

The main task of the vanilla model is to estimate the probability that a student has mastered the knowledge at time step *t*, denoted by a learning parameter $P(L_t)$, $t \geq 0$. The model updates the probability $P(L_t)$ after each opportunity to apply knowledge given an observed correct or incorrect response as follows:

$$P(Correct_t) = \frac{P(L_{t-1})(1-P(S))}{P(L_{t-1})(1-P(S))+(1-P(L_{t-1}))P(G)} \tag{3.1}$$

$$P(Incorrect_t) = \frac{P(L_{t-1})P(S)}{P(L_{t-1})P(S)+(1-P(L_{t-1}))(1-P(G))} \tag{3.2}$$

If $evidence_t \in \{Correct_t, Incorrect_t\}$ means the correctness of a student's answer of the *t-th* opportunity to apply knowledge, the updated probability for the following time step is defined as:

$$P(evidence_t) = P(evidence_t) + \left(1 - P(evidence_t)\right) * P(T) \tag{3.3}$$

At first, the model calculates the probability that the student knew the answer before making an attempt, using the evidence from the current step. Then, taking this into account, it computes the likelihood that the student learned it after the attempt.

Regarding the BKT parameter estimation procedure, Corbett and Anderson [10] discussed individualization per skill and individualization per student of all four BKT parameters. While the individualized BKT model resulted in a better correlation between actual and expected accuracy across students, when compared to the non-individualized BKT model, the accuracy of predicting student test scores (after a period of working with a tutoring system) did not improve tangibly [35]. Finally, the parameter fitting procedure of the vanilla model relates to *expert-based estimations* of the four BKT parameters per skill.

4. METHODOLOGY

The methodology used to review the BKT enhancements and evaluation approaches is in line with the *PRISMA guidelines* [24] consisting of (i) Rationale, objectives and research questions (ii) Eligibility criteria, information sources, and a search strategy (iii), Screening process and study selection and (iv) Data collection and features. Since we have elaborated on the rationale and objectives in the previous sections, we proceed with the criteria, sources and search strategy of works that fall under the scope of this systematic review.

The main eligibility criterion referred to scientific works that aimed to enhance the vanilla BKT model and that were published in the relevant scientific databases until 2022. The implementations of the BKT enhancements could proceed in two directions: the extension of the Bayesian network architecture *and/or* the research of new computational methods in a certain context.

We searched the scientific databases indexing quality-proven journals and conference proceedings, including the Web of Science (Core Collection), Scopus, ACM (Full-Text Collection), IEEE Xplore, and Google Scholar (accessed on March 10, 2023). The search strategy included the phrase *knowledge tracing* and versions of the *Bayes* and *probabilistic* terms that were included in the publication abstracts. Due to the extensiveness of the Google Scholar database, the publication titles were searched for the complete phrase *Bayesian knowledge tracing.* The search details including the number of publication results are presented in Table 4.1.

Table 4.1: Database search details.

Database	**Search query (- 2022)**	**#**
Web of Science (CC)	"knowledge tracing" (Abstract) AND (bayes* OR probab*) (Abstract)	89
Scopus	ABS ("knowledge tracing") AND ABS ((bayes* OR probab*))	200
ACM (Full-Text Collection)	[Abstract: "knowledge tracing"] AND [[Abstract: bayes*] OR [Abstract: probab*]]	25
IEEE Xplore	("Abstract":"knowledge tracing") AND ("Abstract":bayes* OR "Abstract":probab*)	20
Google Scholar	allintitle: "bayesian knowledge tracing"	75

Figure 4.1 shows the PRISMA flow diagram of the publication identification and screening process. Out of 409 results from the five academic databases, we compiled 223 publications (177 duplicates and 9 conference proceedings removed).

The screening of abstracts resulted in the exclusion of 84 publications. They were out of the scope, review papers, available in languages other than English or as a programming code.

In the second phase of screening of 139 full-text manuscripts, we further excluded 83 publications due to the eligibility criteria, not retrieval, or the language other than English.

During the full-text reading phase, the resulting 56 publications were used as a source of additional 17 publications that were included in the review. The additional publications were part of specific events (e.g. Chang et al [36] presented on the 21st Annual meeting of the American Association for Artificial Intelligence), conferences not indexed in scientific databases for the given year (e.g. International Conference on Educational Data Mining in 2014) or works indexed using different keywords (e.g. Baker et al [37]).

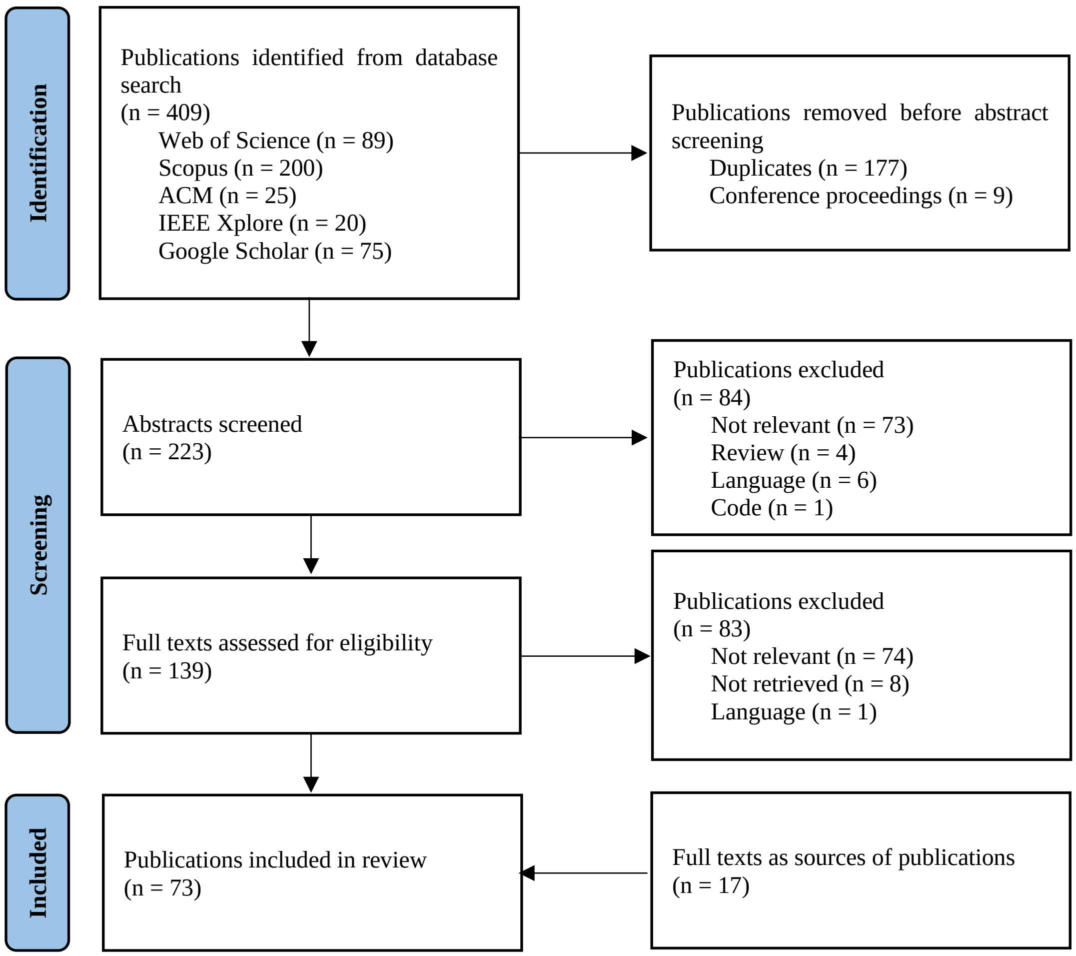

Figure 4.1: PRISMA flow diagram of the publication identification and screening process.

Finally, 73 publications were encompassed in this review study, including the vanilla BKT publication [10].

To get a closer insight into the publications included in the review, we provided the yearly heatmap of the most frequent sources of BKT research in Table 4.2, in which ‘Other’ denotes sources with a single identified publication.

The most common sources were scientific conferences, including the International Conference on Educational Data Mining (IC EDM), International Conference of Intelligent Tutoring Systems (IC ITS), International Conference of Artificial Intelligence in Education (IC AIED), International Conference on User Modelling, Adaptation, and Personalization (IC UMAP), ACM Conference on Learning at Scale (ACMC L@S), International Conference on Learning Analytics & Knowledge (IC LA&K), IEEE Conference on Big Data (IEEEC BigData) and the User Modelling and User-Adapted Interaction (UMUAI) Journal. There was increase in the number of publications in 2008, 2010, and between 2013 and 2018.

Table 4.2: Heatmap of the most frequent sources of publications in the research of BKT enhancements.

Conf./Journal	**1995**	**2004**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2020**	**2021**	**2022**	**T**
IC EDM					1	1	2	1	1	5	3	1	5		3				23
IC ITS		1			2				2		2		2						9
IC AIED				1						2				1			1	1	6
IC UMAP							2	1				1	2						6
ACM C L@S												2	1			1			4
IC LA&K												2	1						3
IEEE BigData														1	1				2
UMUAI	1				1														2
Other			1	1							3	3	1	2	3		2	2	18
Total (T)	1	1	1	2	4	1	4	2	3	7	8	9	12	4	7	1	3	3	**73**

To address the RQ1 and elaborate on various enhancements of the BKT, we found the vanilla model assumptions as appropriate review criteria. The vanilla assumptions derive from the architectural and educational context-based properties of the vanilla BKT model proposed by Corbett and Anderson in [10]. The architectural properties refer to the Hidden Markov model elements, including the nodes with corresponding states and the relationships between nodes (assumptions *A01-A07* in the following text). The educational context-based properties include the vanilla assumptions on the knowledge components dependence, question difficulty and answer attempts (*A08-A10*).

The theory of knowledge inference in the vanilla model consists of the knowledge node with the binary learned and unlearned state (*A01*) and the performance node with the binary correct and incorrect state (*A02*). The prior knowledge, guessing, slipping, and learning parameters are defined as expert-based probabilities estimated per skill (*A03-A06*). The model follows the no-forgetting paradigm by omitting the transition from the learned to the unlearned state (*A07*). The independent knowledge components (*A08*) consist of sets of equally difficult questions used during the knowledge inference process (*A09*). Although a student may have multiple attempts to answer the question in the educational platform, the vanilla model counts only the first attempt (*A10*).

Besides architectural and educational context-based enhancements, to address the RQ1, we reviewed the computational methods used in the enhanced BKT models.

Regarding the RQ2, each publication that proposed enhancements evaluated the approaches by using the datasets from specific educational platforms. Although the diversity and specificity of these studies did not allow the direct comparison of achieved results, this review study provides more insights into the evaluation approaches.

5. BKT ENHANCEMENTS

This study aimed to give the overview of BKT enhancements encompassed by the identified research studies. The identified studies resulted in 62 enhanced BKT models. Some publications referred to the enhancements of the same model, so they were noted as multiple sources of the single enhanced BKT model (e.g. [36], [38]). For more than one source publication per model, we considered the year of the earlier publication as a model source year.

While some of the BKT models addressed the architectural and educational context-based properties of the vanilla BKT model (*A01-A10*), they also proposed new computational methods. On the other side, there were studies focused on the new computational methods without changing the architectural or educational context-based properties. Therefore, we found important to analyse these two aspects independently.

The analysis described in the subsection 5.1. reviews the enhancements of the architectural and educational context-based properties, including the specific enhancements of each of the four BKT parameters (*A03-A06*). The overview of computational methods used for the parameter estimation is given in the subsection 5.2. The evaluation approaches of enhanced BKT models are presented in the following subsection 5.3.

5.1. The architectural and educational context-based enhancements (RQ1)

To summarize the enhanced BKT models, we proposed the enhancement criteria that are in line with the vanilla BKT model assumptions. The enhancement criteria are the result of an iterative analysis of the identified research studies and represent a unique way of classifying the BKT enhancements. Besides the criteria found in the vanilla BKT model (enhancements *E01-E10* in Table 5.1), there were enhancements that extended the vanilla BKT model with new aspects. Additional vanilla BKT enhancements included Student characteristics (*E11*), Tutor interventions (*E12*), and Noise in data (*E13*). Table 5.1 shows the complete list of BKT enhancements and the related vanilla model assumptions.

Table 5.1: Enhancement criteria used to review BKT models.

BKT Enhancements (E)		**Vanilla BKT model Assumptions (A)**	
E01	Knowledge states	A01	Knowledge component node in the Bayesian network includes the binary *learned* and *unlearned* state
E02	Performance states	A02	Performance node in the Bayesian network includes the binary *correct* and *incorrect* state
E03	Prior knowledge	A03	The prior knowledge probability is defined per skill
E04	Guessing	A04	There is a probability of guessing defined per skill
E05	Slipping	A05	There is a probability of slipping defined per skill
E06	Learning	A06	The learning transition probability is defined per skill
E07	Forgetting paradigm	A07	The *no-forgetting paradigm* is followed meaning that there is no transition from learned to unlearned state
E08	Domain knowledge properties	A08	Domain knowledge fractionates into *independent* knowledge components
E09	Question difficulty	A09	Questions of each knowledge component are of *equal difficulty*
E10	Multiple attempts	A10	The *first attempt* to answer the question counts during the modelling process
E11	Student characteristics	Not included	
E12	Tutor interventions	Not included	
E13	Noise in data	Not included	

Although each change in the Bayesian network architecture directly implied the update of BKT parameters, E03-E06 criteria encompassed BKT models with the primary focus on the prior knowledge, guessing, slipping and learning BKT parameters, e.g. Contextual Guess and Slip method [39]. The summarized results of the review of BKT enhancements are presented using a yearly heatmap in Table 5.2. The total of 54 enhanced BKT models addressed the architectural and educational context-based properties of the vanilla BKT model.

Table 5.2: The heatmap of the research on BKT enhancements.

BKT Enhancements (E)		**2004**	**2006**	**2008**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**T**
E01	Knowledge states			1								1	2			1		5
E02	Performance states			1	1					1	2					1		6
E03	Prior knowledge				1			2	1	2	1	1						8
E04	Guessing			1	1	1						1	1					5
E05	Slipping			1	1	2						1	1					6
E06	Learning							3					1					4
E07	Forgetting paradigm		1	1		1				1	1		1					6
E08	Domain know. prop.							1	3	1	3	1		1			2	12
E09	Question difficulty					1		1	3	1	1	1	1					9
E10	Multiple attempts			1	1			1	1	1	1				1			7
E11	Student characteristics			1	1		3		4	1	4		4			1	1	20
E12	Tutor interventions		1	2	2					1	2	1	1					10
E13	Noise in data	1								1							1	3
Total (T)		1	2	9	8	5	3	8	12	10	15	7	12	1	1	3	4	**101**

The first enhanced BKT model emerged in 2004, a decade after the vanilla model. Over the years, the enhancements were continuously investigated and the most frequent additions to the vanilla model included Student characteristics (20 research studies), Domain knowledge properties (12 r.s.), Tutor interventions (10 r.s.), and Question difficulty (9 r.s.). There was decrease in the research after 2018, probably due to the COVID-19 pandemic.

Since each examined research study could enhance one or more of the proposed criteria, we analysed the most frequent variations of the investigated BKT enhancements. It is worth noting that a single criterion represents the simplest enhancement variation. The results are presented in Table 5.3 and variations found in a single research study are summarized as 'Other'.

Table 5.3: The variations of investigated BKT enhancements.

#	**BKT Enhancements (E)**	**# Enhanced BKT models**
1	E11 Student characteristics	9
2	E08 Domain knowledge properties	7
3	E03 Prior knowledge	5
4	E04 Guessing, E05 Slipping, E11 Student characteristics, E12 Tutor interventions	3
5	E09 Question difficulty, E11 Student characteristics	2
6	E11 Student characteristics, E12 Tutor interventions	2
7	E13 Noise in data	2
8-31	Other	24

Among 31 type of enhancement variations, the results indicate that the single criteria of Student characteristics, Prior knowledge and Domain knowledge properties were the most frequently investigated variations. The most frequent combination of enhancements found in 3 research studies included Guessing, Slipping, Student characteristics and Tutor interventions criteria.

The research related to each BKT enhancement is shown in Table 5.4.

Table 5.4: Enhanced BKT models per proposed criteria.

BKT Enhancements (E)		**BKT models**
E01	Knowledge states	Halpern et al., 2018; Liu et al., 2021; Schodde et al., 2017; Yudelson et al., 2008; Zhang & Yao, 2018 [40]–[44]
E02	Performance states	David et al., 2016; Liu et al., 2021; Ostrow et al., 2015; Y. Wang et al., 2010; Y. Wang & Heffernan, 2013; Z. Wang et al., 2016; Yudelson et al., 2008 [41], [43], [45]–[49]
E03	Prior knowledge	Eagle, Corbett, Stamper, McLaren, Baker, et al., 2016; Eagle, Corbett, Stamper, McLaren, Wagner, et al., 2016; Eagle et al., 2017; Nedungadi & Remya, 2014, 2015; Pardos & Heffernan, 2010a; Song et al., 2015; S. Wang et al., 2017; Xu & Mostow, 2013; Yudelson et al., 2013 [35], [50]–[58]
E04	Guessing	Agarwal et al., 2018; Baker et al., 2008b, 2008a, 2010; Pardos & Heffernan, 2011; Zhou et al., 2017 [37], [39], [59]–[62]
E05	Slipping	Agarwal et al., 2018; Baker et al., 2008b, 2008a, 2010; Pardos & Heffernan, 2011; Qiu et al., 2011; Zhou et al., 2017 [37], [39], [59]–[63]
E06	Learning	Adjei et al., 2013; Baker et al., 2018; Sao Pedro et al., 2013; Yudelson et al., 2013 [35], [64]–[66]
E07	Forgetting paradigm	Beck et al., 2008; Chang et al., 2006b; Halpern et al., 2018; Khajah et al., 2016; Nedungadi & Remya, 2015; Qiu et al., 2011; Yudelson et al., 2008 [31], [36], [38], [40], [43], [54], [63]
E08	Domain know. prop.	González-Brenes et al., 2014; Huang et al., 2016; Huang & Brusilovsky, 2016; Khajah et al., 2016; Meng et al., 2019; Sao Pedro et al., 2013, 2014; Z. Wang et al., 2016; Hawkins & Heffernan, 2014; MacHardy, 2015; MacHardy & Pardos, 2015; Z. Wang et al., 2016; Sun et al, 2022; Chan et al, 2022 [31], [49], [49], [66]–[76]
E09	Question difficulty	Baker et al., 2018; David et al., 2016; González-Brenes et al., 2014; Khajah, Huang, et al., 2014; Khajah, Wing, et al., 2014; Ostrow et al., 2015; Pardos et al., 2013; Pardos & Heffernan, 2011; Zhou et al., 2017 [45], [46], [61], [62], [65], [67], [77]–[79]
E10	Multiple attempts	Bhatt et al., 2020; Gonzalez-Brenes et al., 2014; Pardos et al., 2013; Yudelson et al., 2008 [43], [67], [79], [80]
E11	Student characteristics	Agarwal et al., 2018; Baker et al., 2008b, 2008a, 2010; Eagle et al., 2018; Khajah, Wing, et al., 2014; Khajah et al., 2016; Nedungadi & Remya, 2014; Yudelson, 2021; Zhu et al., 2018; Xu et al., 2014; Corrigan et al., 2015; Spaulding et al., 2016; Rau & Pardos, 2016; Lin et al., 2016; Lin & Chi, 2016; Halpern et al., 2018; Khajah, Huang, et al., 2014; J. I. Lee & Brunskill, 2012; Pardos et al., 2012; Y. Wang & Heffernan, 2012; Gorgun & Bulut, 2022 [31], [37], [39], [40], [53], [59], [60], [77], [78], [81]–[93]
E12	Tutor interventions	Agarwal et al., 2018; Baker et al., 2008b, 2008a, 2010; Beck et al., 2008; Chang et al., 2006b; Lin et al., 2016; Lin & Chi, 2016; Ostrow et al., 2015; Rau & Pardos, 2016; Schodde et al., 2017; Y. Wang et al., 2010; Y. Wang & Heffernan, 2013; Yudelson et al., 2008 [36]–[39], [42], [43], [46]–[48], [59], [60], [87]–[89]
E13	Noise in data	Beck & Sison, 2004; Falakmasir et al., 2015; Gorgun & Bulut, 2022 [93]–[95]

5.2. Computational methods (RQ1)

Regarding the computational methods used in the proposed BKT approaches, they were generally related to the estimation of BKT parameters. As we already mentioned, there were models that did not have to interfere with the vanilla model assumptions, but primarily addressed the computational challenges, e.g. Dirichlet priors method [96], [97]. Overall, 56 enhanced BKT models reported the use of computational methods.

The results of the review of computational methods used in the research of BKT enhancements found in over 2 identified research studies are presented in Table 5.5.

Table 5.5: Computational methods used in the research of BKT enhancements.

#	Computational methods with over 2 applications	# Enhanced BKT models
1	Expectation-Maximization method	24
2	Markov Chain Monte Carlo (MCMC) method	5
3	Brute force method	4
4	K-means clustering	4
5	Contextual Guess and Slip method	3
6	Knowledge Heuristics and Empirical probabilities method	3
7-31	Other	32

In terms of the computational methods, most of the research aimed to improve the expert-based estimations of BKT parameters used in the vanilla model. In that sense, the Expectation-Maximization method, firstly used in 2006, practically became the standard (24 research studies). The other computational methods included the Monte Carlo method (5 r.s.), the Brute force method (4 r.s.), K-means clustering (4 r.s.), the Contextual Guess and Slip method (3 r.s.), and the Knowledge Heuristics with Empirical Probabilities method (3 r.s.).

The increased interest in the research of the Expectation-Maximization method resulted in detection of difficulties. Although firstly reported as the model identifiability problem [97], Doroudi and Brunskill [98] revealed that under mild conditions on the parameters, the BKT model is actually identifiable and it 'only' suffers from the local optima problem. The recent open-source accessible and computationally efficient Python library of BKT models – pyBKT [30] also includes the Expectation-Maximization method to fit the BKT parameters. Since it is well known that the Expectation-Maximization algorithm is susceptible to converging to local optima of the likelihood function rather than converging to the global optimum (local optima problem), the pyBKT runs multiple iterations of the algorithm with different initializations of the parameters to avoid this problem [30]. Besides, it was reported that the Expectation-Maximization method is prone to the semantic model degeneracy, meaning that it can be inconsistent with the conceptual assumptions underlying the BKT model [98].

5.3. Evaluation approaches (RQ2)

In terms of the evaluation approaches, we reviewed educational platforms and performance measures included in the research of BKT enhancements.

Table 5.6 shows a yearly heatmap of the educational platforms used in the reviewed publications. We summarized those platforms with a single application as 'Other'.

Besides ITSs, we found the application of the BKT model enhancements in MOOCs, game-based platforms, and online learning platforms in the field of human resources. The research on the BKT enhancements typically included the ASSISTments (19 r.s.) and the Cognitive Tutor (19 research studies). Other educational platforms with over 2 applications referred to Massive Open Online Courses (5 r.s.) and simulated datasets (7 r.s.). The MOOC environments included the edX [79], the Coursera [49], the Khan Academy [31], [74], and the Junyi Academy [83].

Table 5.6: Educational platforms used in the research of BKT enhancements.

Educational platforms	**2004**	**2006**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**T**
Assistments					2	2	2	1	3	3	1	1	1			1	2	19
Cognitive Tutor			1	1	1	2	1	2	2	2	2		1	1		2	1	19
Simulated data					1			2			1		2	1				7
MOOC								1		1	2		1					5
Andes Tutor									2									2
Inq-ITS								1	1									2
JavaGuide									1		1							2
Reading Tutor		1							1									2
Robot Tutor											1	1						2
Other	1		1						2	1	6	2	3	1	1			18
Total (T)	1	1	2	1	4	4	3	7	12	7	14	4	8	3	1	3	3	**78**

As for the domain, the examined datasets were related to Math (38 r.s.), Language learning and Programming (per 6 r.s.), Genetics, Physics and Engineering (per 3 r.s.), Science (per 2 r.s.), and Medicine and Chemistry (per single r.s.).

The only publicly available datasets identified in the review were related to the environments of the Cognitive Tutor and the ASSISTments. These datasets were part of the Educational Data Mining KDD Cup Challenge, hosted by PSLC DataShop [99], that highly contributed to the awareness of the importance of replicability and comparison of the proposed models.

Regarding the used performance measures, Table 5.7 shows the most frequently used measures in the research of BKT enhancements with over 2 applications.

Table 5.7: Performance measures used in the research of BKT enhancements.

#	**Performance measures with over 2 applications**	**# Enhanced BKT models**
1	RMSE	27
2	AUC-ROC	22
3	Accuracy	19
4	MAE	8
5	Correlation	3
6-14	Other	13

The most frequently used performance measures included the RMSE measure (Root Mean Square Error, 27 r.s.), the AUC-ROC (Area Under Curve, Receiver Operating Characteristics curve, 22 r.s.) and the Accuracy measure (19 r.s.). These performance measures are frequently used metrics for classification tasks in the machine learning field.

Overall, the research of BKT enhancements included two types of model evaluations, including (i) the prediction of in-tutor performance as the correctness of the following student's answer and (ii) the ability to estimate overall knowledge mastery. The first type of evaluation was frequently applied in the literature and used performance measures such as RMSE, AUC-ROC and Accuracy. In terms of the knowledge mastery prediction, only a few research studies investigated the relationship between knowledge estimated by the system and knowledge demonstrated on the post-test outside of the system's environment [42], [43], [60], [88], [89], [94].

The researchers typically compared the proposed models to the vanilla BKT model and despite the variety of BKT models over the years, the accessible and easy to use BKT implementations remained elusive. There were only three available BKT implementation frameworks, including the Bayes Net Toolbox [100], [101], the hmm-scalable implementation [35] and the approach proposed by Xu et al. [102] for MOOC resources. While most of the enhanced BKT models reported better results than the vanilla model, the unexpected and mixed results for both types of evaluation approaches were reported in the literature.

The unexpected performance results as the model drawbacks were reported in the case of the Help BKT model [38], [100], the Affective BKT model [85], the BKT with eye-tracking model [87] and the model by Adjei et al [64]. The parameters of the Help BKT model suggested that students benefited from the scaffolding and teaching effects of help. Despite that information, the Help BKT model did not outperform the vanilla model. The possible reason for such results may be the overfitting of the data. In the case of the Affective BKT, the model did not offer additional predictive power beyond vanilla model, probably due to the lack of variability in the binary affective state measured across student responses. The BKT with eye-tracking data did not add information relevant to students' representation skills, possibly for the same reason as the Affective BKT. Adjei et al concluded that the different learning rates based on the answer correctness did not lead to better model predictions.

In terms of the mixed performance results, several research studies reported ambiguous results dependent on the educational settings, including the BKT with Contextual guess and slip method (CGS-BKT) [37], [39], the Item Difficulty Effect Model (KT-IDEM) [61], the Student Skill BKT model (SS-BKT) [92] and the BKT with tutoring actions [42]. While the CGS-BKT showed better results than the vanilla model in predicting in-tutor performance, the model performed much more poorly on the post-test. In the case of the KT-IDEM, the model provided reliably better in-tutor performance prediction on the ASSISTments dataset but was not significantly different from vanilla model in the case of the Cognitive Tutor. Also, it was found that the SS-BKT model was investigated under the simulated conditions and outperformed the vanilla model only when the number of students and skills were large. As in the case of the CGS-BKT, the BKT with tutoring actions did not show a significant difference in the post-test results.

6. CONCLUSION

It has been 25 years of BKT research, and the vanilla model is still a representative Bayesian network-based approach. Over the years, various improvements have been proposed in the literature, mostly outperforming the vanilla model, but their limited availability negatively affected their further application. Even the latest research on deep learning-based knowledge tracing revealed that just enabling the forgetting parameter in the vanilla model led to similar results as the neural network-based model.

Because of the specificities of the educational platforms and subsets of data used to train the models, there is no possibility to compare the achieved performance results of the proposed enhanced models. Moreover, there has been no systematic review of the BKT enhancements since its introduction, as the existing research reviewed only subsets of enhanced models. The most extensive student modelling review focused on ML approaches and encompassed 18 student models based on Bayesian networks [11]. The other review study elaborated on the 8 BKT models from 2010 to 2015 [15]. The systematic review from 2013 discussed the 13 ML student models and 18 Bayesian network-based models [19]. In addition, each publication that proposes enhanced BKT model also provides only subset of the background literature because of the limited format and different purpose. Finally, this study brings a systematic and more exhaustive review by encompassing 62 BKT models that aimed to enhance the vanilla model.

To summarize the research on the BKT enhancements, we proposed the unique set of criteria, including 10 aspects based on the vanilla model assumptions and 3 aspects new to the vanilla model. The most frequently improved aspects were additional to the vanilla model and included student characteristics and tutor interventions. The other frequently investigated aspects included the domain knowledge properties (assumed as independent knowledge components in the vanilla model) and the question difficulty (assumed as constant in the vanilla model). Although less investigated in the literature, the obvious drawbacks of the vanilla model referred to the binary states and exclusion (common prior knowledge and only the first answer attempt considered in the vanilla model). As suggested in the reference literature, the uncertainty and fuzziness in this context would be more appropriate. Besides, knowledge does fade, and as research shows, it occurs in such a short period that the inclusion of the forgetting parameter is mandatory.

Overall, the BKT enhancements can be differed as generally applicable and dependent on the capabilities of educational settings. While most of the enhancements already defined by the vanilla model represent generally applicable enhancements, if the educational environment enables specific features (e.g., student characteristics, tutor interventions), the research suggests checking their contribution to the knowledge inference process.

Besides the previous enhancement criteria, the BKT models were reviewed according to the incorporated computational methods. The enhanced models generally improved the expert-based estimations of BKT parameters assumed by the vanilla model. The early introduction of the Expectation-Maximization method proved efficient and has become the standard in this context. However, we expect that the novel ML methods will eventually contribute to the research challenges related to the BKT.

The major limitation in the BKT research is model availability. The positive example in this context is the pyBKT library [30], a recently introduced accessible and computationally efficient BKT implementation framework. Besides general features, the generalized BKT model requires an accessible and easy to use implementation framework.

Regarding the future research, various educational environments represent a broad testing ground and we find interesting to investigate the potential application of BKT models in the context of the Moodle LMS and the higher education course with over 150 students. Based on the BKT research, adaptive and individualized formative tests will be introduced to facilitate higher learning engagement and performance. In this sense, we seek for the student model that will enable the prediction of the overall student performance based on the weekly estimations of knowledge mastery.

BIBLIOGRAPHY

[1] B. F. Skinner, “The science of learning and the art of teaching,” *Harvard Educational Review*, vol. 24, pp. 86–97, 1954.

[2] D. Sleeman and J. S. Brown, “Introduction: Intelligent Tutoring Systems: An Overview,” in *Intelligent Tutoring Systems*, Sleeman, D.H., Brown, J.S., Academic Press, Burlington, MA, 1982, pp. 1–11.

[3] E. Wenger, *Artificial Intelligence and Tutoring Systems*. Morgan Kaufmann Publishers, Inc., California, USA, 1987.

[4] J. A. DeFalco and A. M. Sinatra, “Adaptive Instructional Systems: The Evolution of Hybrid Cognitive Tools and Tutoring Systems,” in *Adaptive Instructional Systems*, Cham, 2019, pp. 52–61. doi: 10.1007/978-3-030-22341-0_5.

[5] R. Freedman, S. S. Ali, and S. McRoy, “Links: What is an Intelligent Tutoring System?,” *Intelligence*, vol. 11, no. 3, pp. 15–16, Sep. 2000, doi: 10.1145/350752.350756.

[6] R. Nkambou, J. Bourdeau, and R. Mizoguchi, Eds., *Advances in Intelligent Tutoring Systems*, vol. 308. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. doi: 10.1007/978-3-642-14363-2.

[7] H. S. Nwana, “Intelligent tutoring systems: an overview,” *Artif Intell Rev*, vol. 4, no. 4, pp. 251–277, Dec. 1990, doi: 10.1007/BF00168958.

[8] J. A. Self, “Student models in computer-aided instruction,” *International Journal of Man-Machine Studies*, vol. 6, no. 2, pp. 261–276, 1974.

[9] B. P. Woolf, “AI in Education,” in *Artificial Intelligence – Encyclopedias*, John Willy & Sons, 1992, pp. 434–444.

[10] A. T. Corbett and J. R. Anderson, “Knowledge tracing: Modeling the acquisition of procedural knowledge,” *User Modeling and User-Adapted Interaction*, vol. 4, no. 4, pp. 253–278, 1995.

[11] Q. Liu, S. Shen, Z. Huang, E. Chen, and Y. Zheng, “A Survey of Knowledge Tracing,” *arXiv:2105.15106 [cs]*, 2021, Accessed: Aug. 24, 2021. [Online]. Available: http://arxiv.org/abs/2105.15106

[12] S. I. Ramírez Luelmo, N. El Mawas, and J. Heutte, “Machine Learning Techniques for Knowledge Tracing: A Systematic Literature Review:,” in *Proceedings of the 13th International Conference on Computer Supported Education*, Online Streaming, --- Select a Country ---, 2021, pp. 60–70. doi: 10.5220/0010515500600070.

[13] R. Pelánek, “Bayesian Knowledge Tracing, Logistic Models, and Beyond: An Overview of Learner Modeling Techniques,” *User Modeling and User-Adapted Interaction*, vol. 27, no. 3–5, pp. 313–350, Dec. 2017, doi: 10.1007/s11257-017-9193-2.

[14] M. Anouar Tadlaoui, A. Souhaib, M. Khaldi, and R. Carvalho, “Learner Modeling in Adaptive Educational Systems: A Comparative Study,” *International Journal of Modern Education and Computer Science*, vol. 8, pp. 1–10, 2016, doi: 10.5815/ijmecs.2016.03.01.

[15] S. M. Sani, A. B. Bichi, and S. Ayuba, “Artificial intelligence approaches in student modeling: half decade review (2010-2015),” *International Journal of Computer Science and Network*, vol. 5, no. 5, Art. no. 5, 2016.

[16] L. D. Kurup, A. Joshi, and N. Shekhokar, “A review on student modeling approaches in ITS,” in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2016, pp. 2513–2517.

[17] A. Zafar and N. Ahmad, “An overview of Student Modeling Approaches under Uncertain Conditions,” *International Journal of Information Technology and Management*, vol. 4, no. 1, 2013.

[18] P. Pavlik, K. Brawner, A. Olney, and A. Mitrovic, “A Review of Learner Models Used in Intelligent Tutoring Systems,” in *Design Recommendations for Intelligent Tutoring Systems - Learner Modeling*, vol. Volume I, Army Research Labs, University of Memphis, 2013, pp. 39–68.

[19] K. Chrysafiadi and M. Virvou, “Student modeling approaches: A literature review for the last decade,” *Expert Systems with Applications*, vol. 40, no. 11, pp. 4715–4729, Sep. 2013, doi: 10.1016/j.eswa.2013.02.007.

[20] B. Harrison and D. Roberts, “A Review of Student Modeling Techniques in Intelligent Tutoring Systems,” in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 2012, vol. 8, pp. 61–66. Accessed: Jul. 25, 2021. [Online]. Available: https://ojs.aaai.org/index.php/AIIDE/article/view/12574

[21] M. C. Desmarais and R. S. Baker, “A review of recent advances in learner and skill modeling in intelligent learning environments,” *User Model User-Adap Inter*, vol. 22, no. 1, pp. 9–38, Apr. 2012, doi: 10.1007/s11257-011-9106-8.

[22] M. Vandewaetere, P. Desmet, and G. Clarebout, “The contribution of learner characteristics in the development of computer-based adaptive learning environments,” *Computers in Human Behavior*, vol. 27, no. 1, pp. 118–130, Jan. 2011, doi: 10.1016/j.chb.2010.07.038.

[23] P. Brusilovsky and E. Millán, “User Models for Adaptive Hypermedia and Adaptive Educational Systems,” in *The Adaptive Web: Methods and Strategies of Web Personalization*, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds. Berlin, Heidelberg: Springer, 2007, pp. 3–53. doi: 10.1007/978-3-540-72079-9_1.

[24] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and T. P. Group, “Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement,” *PLOS Medicine*, vol. 6, no. 7, p. e1000097, 2009, doi: 10.1371/journal.pmed.1000097.

[25] T. M. Mitchell, *Machine Learning*. New York: McGraw-Hill Education, 1997.

[26] H. Cen, K. Koedinger, and B. Junker, “Learning Factors Analysis – A General Method for Cognitive Model Evaluation and Improvement,” in *Intelligent Tutoring Systems*, Berlin, Heidelberg, 2006, pp. 164–175. doi: 10.1007/11774303_17.

[27] P. I. Pavlik, H. Cen, and K. R. Koedinger, “Performance Factors Analysis --A New Alternative to Knowledge Tracing,” in *Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling*, NLD, Jul. 2009, pp. 531–538.

[28] C. Piech *et al.*, “Deep Knowledge Tracing,” in *Advances in Neural Information Processing Systems*, 2015, vol. 28. Accessed: Feb. 19, 2022. [Online]. Available: https://papers.nips.cc/paper/2015/hash/bac9162b47c56fc8a4d2a519803d51b3-Abstract.html

[29] S. Montero, A. Arora, S. Kelly, B. Milne, and M. Mozer, “Does deep knowledge tracing model interactions among skills?,” presented at the The 11th International Conference on Educational Data Mining, EDM 2018, 2018. Accessed: Oct. 06, 2021. [Online]. Available: http://www.scopus.com/inward/record.url?scp=85084011386&partnerID=8YFLogxK

[30] A. Badrinath, F. Wang, and Z. A. Pardos, “pyBKT: An Accessible Python Library of Bayesian Knowledge Tracing Models,” *CoRR*, vol. abs/2105.00385, 2021, Accessed: Oct. 06, 2021. [Online]. Available: https://arxiv.org/abs/2105.00385

[31] M. Khajah, R. V. Lindsey, and M. Mozer, “How Deep is Knowledge Tracing?,” in *Proceedings of the 9th International Conference on Educational Data Mining, EDM 2016, Raleigh, North Carolina, USA, June 29 - July 2, 2016*, Raleigh, North Carolina, USA, 2016. Accessed: Dec. 03, 2021. [Online]. Available: http://www.educationaldatamining.org/EDM2016/proceedings/paper_144.pdf

[32] R. S. Baker and P. S. Inventado, “Educational Data Mining and Learning Analytics,” in *Learning Analytics: From Research to Practice*, J. A. Larusson and B. White, Eds. New York, NY: Springer, 2014, pp. 61–75. doi: 10.1007/978-1-4614-3305-7_4.

[33] J. R. Anderson, *Rules of the mind*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc, 1993, pp. ix, 320.

[34] R. C. Atkinson, “Optimizing the learning of a second-language vocabulary,” *Journal of Experimental Psychology*, vol. 96, no. 1, pp. 124–129, 1972, doi: 10.1037/h0033475.

[35] M. Yudelson, K. R. Koedinger, and G. J. Gordon, “Individualized Bayesian Knowledge Tracing Models,” in *Artificial Intelligence in Education*, 2013, pp. 171–180.

[36] K. Chang, J. E. Beck, J. Mostow, and A. T. Corbett, “Does Help Help? A Bayes Net Approach to Modeling Tutor Interventions,” presented at the AAAI2006 Workshop on Educational Data Mining, Boston, Massachusetts, 2006.
[37] R. S. Baker, A. T. Corbett, and V. Aleven, “Improving Contextual Models of Guessing and Slipping with a Truncated Training Set,” presented at the International Conference on Educational Data Mining, EDM 2008, 2008. doi: 10.1184/R1/6470135.v1.
[38] J. E. Beck, K. Chang, J. Mostow, and A. T. Corbett, “Does Help Help? Introducing the Bayesian Evaluation and Assessment Methodology,” in *Intelligent Tutoring Systems, 9th International Conference, ITS 2008, Montreal, Canada, June 23-27, 2008, Proceedings*, Montreal, Canada, 2008, vol. 5091, pp. 383–394. doi: 10.1007/978-3-540-69132-7_42.
[39] R. S. Baker, A. T. Corbett, and V. Aleven, “More Accurate Student Modeling through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing,” in *Proceedings of the 9th Inernational Conference on Intelligent Tutoring Systems, ITS 2008, Montreal, Canada, June 23-27, 2008*, Montreal, Canada, 2008, vol. 5091, pp. 406–415. doi: 10.1007/978-3-540-69132-7_44.
[40] D. Halpern *et al.*, “Knowledge Tracing Using the Brain,” in *Proceedings of the 11th International Conference on Educational Data Mining, EDM 2018, Buffalo, NY, USA, July 15-18, 2018*, Buffalo, NY, USA, 2018. Accessed: Dec. 03, 2021. [Online]. Available: http://educationaldatamining.org/files/conferences/EDM2018/papers/EDM2018_paper_158.pdf
[41] F. Liu, X. Hu, C. Bu, and K. Yu, “Fuzzy Bayesian Knowledge Tracing,” *IEEE Transactions on Fuzzy Systems*, pp. 1–1, 2021, doi: 10.1109/TFUZZ.2021.3083177.
[42] T. Schodde, K. Bergmann, and S. Kopp, “Adaptive Robot Language Tutoring Based on Bayesian Knowledge Tracing and Predictive Decision-Making,” in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, HRI 2017, Vienna, Austria, March 6-9, 2017*, Vienna, Austria, 2017, pp. 128–136. doi: 10.1145/2909824.3020222.
[43] M. Yudelson, O. Medvedeva, and R. S. Crowley, “A multifactor approach to student model evaluation,” *User Modeling and User-Adapted Interaction*, vol. 18, no. 4, pp. 349–382, 2008, doi: 10.1007/s11257-007-9046-5.
[44] K. Zhang and Y. Yao, “A three learning states Bayesian knowledge tracing model,” *Knowledge Based Systems*, vol. 148, pp. 189–201, 2018, doi: 10.1016/j.knosys.2018.03.001.
[45] Y. B. David, A. Segal, and Y. Gal, “Sequencing educational content in classrooms using Bayesian knowledge tracing,” in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, LAK 2016, Edinburgh, United Kingdom, April 25-29, 2016*, Edinburgh, United Kingdom, 2016, pp. 354–363. doi: 10.1145/2883851.2883885.
[46] K. Ostrow, C. Donnelly, S. Adjei, and N. T. Heffernan, “Improving Student Modeling Through Partial Credit and Problem Difficulty,” in *Proceedings of the Second ACM Conference on Learning @ Scale, L@S 2015, Vancouver, BC, Canada, March 14 - 18, 2015*, Vancouver, BC, Canada, 2015, pp. 11–20. doi: 10.1145/2724660.2724667.
[47] Y. Wang, N. T. Heffernan, and J. E. Beck, “Representing Student Performance with Partial Credit,” in *Educational Data Mining 2010, The 3rd International Conference on Educational Data Mining, Pittsburgh, PA, USA, June 11-13, 2010. Proceedings*, Pittsburgh, PA, USA, 2010, pp. 335–336. Accessed: Dec. 03, 2021. [Online]. Available: http://educationaldatamining.org/EDM2010/uploads/proc/edm2010_submission_86.pdf
[48] Y. Wang and N. T. Heffernan, “Extending Knowledge Tracing to Allow Partial Credit: Using Continuous versus Binary Nodes,” in *Artificial Intelligence in Education - 16th International Conference, AIED 2013, Memphis, TN, USA, July 9-13, 2013. Proceedings*, Memphis, TN, USA, 2013, vol. 7926, pp. 181–188. doi: 10.1007/978-3-642-39112-5_19.
[49] Z. Wang, J. Zhu, X. Li, Z. Hu, and M. Zhang, “Structured Knowledge Tracing Models for Student Assessment on Coursera,” in *Proceedings of the Third ACM Conference on Learning @ Scale, L@S 2016, Edinburgh, Scotland, UK, April 25 - 26, 2016*, Edinburgh, Scotland, UK, 2016, pp. 209–212. doi: 10.1145/2876034.2893416.

[50] M. Eagle *et al.*, “Predicting Individual Differences for Learner Modeling in Intelligent Tutors from Previous Learner Activities,” in *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization, UMAP 2016, Halifax, NS, Canada, July 13 - 17, 2016*, Halifax, NS, Canada, 2016, pp. 55–63. doi: 10.1145/2930238.2930255.
[51] M. Eagle *et al.*, “Estimating Individual Differences for Student Modeling in Intelligent Tutors from Reading and Pretest Data,” in *Intelligent Tutoring Systems - 13th International Conference, ITS 2016, Zagreb, Croatia, June 7-10, 2016. Proceedings*, Zagreb, Croatia, 2016, vol. 9684, pp. 133–143. doi: 10.1007/978-3-319-39583-8_13.
[52] M. Eagle *et al.*, “Exploring Learner Model Differences Between Students,” in *Artificial Intelligence in Education - 18th International Conference, AIED 2017, Wuhan, China, June 28 - July 1, 2017, Proceedings*, Wuhan, China, 2017, vol. 10331, pp. 494–497. doi: 10.1007/978-3-319-61425-0_48.
[53] P. Nedungadi and M. S. Remya, “Predicting students’ performance on intelligent tutoring system - Personalized clustered BKT (PC-BKT) model,” in *IEEE Frontiers in Education Conference, FIE 2014, Proceedings, Madrid, Spain, October 22-25, 2014*, Madrid, Spain, 2014, pp. 1–6. doi: 10.1109/FIE.2014.7044200.
[54] P. Nedungadi and M. S. Remya, “Incorporating forgetting in the Personalized, Clustered, Bayesian Knowledge Tracing (PC-BKT) model,” in *2015 International Conference on Cognitive Computing and Information Processing(CCIP)*, 2015, pp. 1–5. doi: 10.1109/CCIP.2015.7100688.
[55] Z. A. Pardos and N. T. Heffernan, “Modeling Individualization in a Bayesian Networks Implementation of Knowledge Tracing,” in *User Modeling, Adaptation, and Personalization*, 2010, pp. 255–266.
[56] Y. Song, Y. Jin, X. Zheng, H. Han, Y. Zhong, and X. Zhao, “PSFK: A Student Performance Prediction Scheme for First-Encounter Knowledge in ITS,” in *Knowledge Science, Engineering and Management - 8th International Conference, KSEM 2015, Chongqing, China, October 28-30, 2015, Proceedings*, Chongqing, China, 2015, vol. 9403, pp. 639–650. doi: 10.1007/978-3-319-25159-2_58.
[57] S. Wang, Y. Han, W. Wu, and Z. Hu, “Modeling student learning outcomes in studying programming language course,” in *2017 Seventh International Conference on Information Science and Technology (ICIST)*, 2017, pp. 263–270. doi: 10.1109/ICIST.2017.7926768.
[58] Y. Xu and J. Mostow, “Using Item Response Theory to Refine Knowledge Tracing,” in *Proceedings of the 6th International Conference on Educational Data Mining, Memphis, Tennessee, USA, July 6-9, 2013*, 2013, pp. 356–357. Accessed: Sep. 05, 2022. [Online]. Available: http://www.educationaldatamining.org/EDM2013/papers/rn_paper_79.pdf
[59] D. Agarwal, N. Babel, and R. S. Baker, “Contextual Derivation of Stable BKT Parameters for Analysing Content Efficacy,” in *Proceedings of the 11th International Conference on Educational Data Mining, EDM 2018, Buffalo, NY, USA, July 15-18, 2018*, Buffalo, NY, USA, 2018. Accessed: Dec. 03, 2021. [Online]. Available: http://educationaldatamining.org/files/conferences/EDM2018/papers/EDM2018_paper_14.pdf
[60] R. S. Baker *et al.*, “Contextual Slip and Prediction of Student Performance after Use of an Intelligent Tutor,” in *User Modeling, Adaptation, and Personalization, 18th International Conference, UMAP 2010, Big Island, HI, USA, June 20-24, 2010. Proceedings*, Big Island, HI, USA, 2010, vol. 6075, pp. 52–63. doi: 10.1007/978-3-642-13470-8_7.
[61] Z. A. Pardos and N. T. Heffernan, “KT-IDEM: Introducing Item Difficulty to the Knowledge Tracing Model,” in *User Modeling, Adaption and Personalization - 19th International Conference, UMAP 2011, Girona, Spain, July 11-15, 2011. Proceedings*, Girona, Spain, 2011, vol. 6787, pp. 243–254. doi: 10.1007/978-3-642-22362-4_21.
[62] X. Zhou, W. Wu, and Y. Han, “Modeling multiple subskills by extending knowledge tracing model using logistic regression,” in *2017 IEEE International Conference on Big Data (Big Data)*, Dec. 2017, pp. 3994–4003. doi: 10.1109/BigData.2017.8258413.
[63] Y. Qiu, Y. Qi, H. Lu, Z. A. Pardos, and N. T. Heffernan, “Does Time Matter? Modeling the Effect of Time with Bayesian Knowledge Tracing,” in *EDM*, 2011.

[64] S. Adjei, S. Salehizadeh, Y. Wang, and N. T. Heffernan, "Do students really learn an equal amount independent of whether they get an item correct or wrong?," in *Proceedings of the 6th International Conference on Educational Data Mining, Memphis, Tennessee, USA, July 6-9, 2013*, 2013, pp. 304–305. Accessed: Sep. 05, 2022. [Online]. Available: http://www.educationaldatamining.org/EDM2013/papers/rn_paper_53.pdf

[65] R. S. Baker, S. M. Gowda, and E. Salamin, "Modeling the Learning That Takes Place Between Online Assessments," in *Proceedings of the 26th International Conference on Computers in Education, Philippines: Asia-Pacific Society for Computers in EducationAsia-Pacific Society for Computers in Education*, Philippines, 2018, p. 8.

[66] M. Sao Pedro, R. S. Baker, and J. D. Gobert, "Incorporating Scaffolding and Tutor Context into Bayesian Knowledge Tracing to Predict Inquiry Skill Acquisition," in *Proceedings of the 6th International Conference on Educational Data Mining, Memphis, Tennessee, USA, July 6-9, 2013*, Memphis, Tennessee, USA, 2013, pp. 185–192. Accessed: Dec. 03, 2021. [Online]. Available: http://www.educationaldatamining.org/EDM2013/papers/rn_paper_27.pdf

[67] J. P. González-Brenes, Y. Huang, and P. Brusilovsky, "General Features in Knowledge Tracing to Model Multiple Subskills, Temporal Item Response Theory, and Expert Knowledge," in *Proceedings of the 7th International Conference on Educational Data Mining, EDM 2014, London, UK, July 4-7, 2014*, London, UK, 2014, pp. 84–91. Accessed: Dec. 03, 2021. [Online]. Available: http://www.educationaldatamining.org/EDM2014/uploads/procs2014/longpapers/84_EDM-2014-Full.pdf

[68] Y. Huang, J. Guerra, and P. Brusilovsky, "Modeling Skill Combination Patterns for Deeper Knowledge Tracing," presented at the The 6th Intl. Workshop on Personalization Approaches in Learning Environments (PALE 2016) in the 24th Conf. on User Modeling, Adaptation and Personalization (UMAP 2016), 2016.

[69] Y. Huang and P. Brusilovsky, "Towards Modeling Chunks in a Knowledge Tracing Framework for Students' Deep Learning," in *Proceedings of the 9th International Conference on Educational Data Mining, EDM 2016, Raleigh, North Carolina, USA, June 29 - July 2, 2016*, 2016, pp. 666–668. Accessed: Sep. 05, 2022. [Online]. Available: http://www.educationaldatamining.org/EDM2016/proceedings/paper_196.pdf

[70] L. Meng, M. Zhang, W. Zhang, and Y. Chu, "CS-BKT: introducing item relationship to the Bayesian knowledge tracing model," *Interactive Learning Environments*, pp. 1–11, Jun. 2019, doi: 10.1080/10494820.2019.1629600.

[71] M. Sao Pedro, Y. Jiang, L. Paquette, R. S. Baker, and J. D. Gobert, "Identifying Transfer of Inquiry Skills across Physical Science Simulations using Educational Data Mining," in *Learning and Becoming in Practice: Proceedings of the 11th International Conference of the Learning Sciences, ICLS 2014, Boulder, Colorado, USA, June 23-27, 2014*, Boulder, Colorado, USA, 2014. Accessed: Dec. 03, 2021. [Online]. Available: https://repository.isls.org/handle/1/1116

[72] W. J. Hawkins and N. T. Heffernan, "Using Similarity to the Previous Problem to Improve Bayesian Knowledge Tracing," in *Proceedings of the Workshops held at Educational Data Mining 2014, co-located with 7th International Conference on Educational Data Mining (EDM 2014), London, United Kingdom, July 4-7, 2014*, London, UK, 2014, vol. 1183. Accessed: Dec. 03, 2021. [Online]. Available: http://ceur-ws.org/Vol-1183/bkt20y_paper04.pdf

[73] Z. MacHardy, "Applications of bayesian knowledge tracing to the curation of educational videos," University of California at Berkeley, 2015.

[74] Z. MacHardy and Z. A. Pardos, "Toward the Evaluation of Educational Videos using Bayesian Knowledge Tracing and Big Data," in *Proceedings of the Second ACM Conference on Learning @ Scale, L@S 2015, Vancouver, BC, Canada, March 14 - 18, 2015*, 2015, pp. 347–350. doi: 10.1145/2724660.2728690.

[75] S. Sun, X. Hu, C. Bu, F. Liu, Y. Zhang, and W. Luo, "Genetic Algorithm for Bayesian Knowledge Tracing: A Practical Application," in *Advances in Swarm Intelligence*, Cham, 2022, pp. 282–293. doi: 10.1007/978-3-031-09677-8_24.

[76] K. I. Chan, R. Tse, and P. I. S. Lei, “Tracing Students’ Learning Performance on Multiple Skills using Bayesian Methods,” in *Proceedings of the 6th International Conference on Education and Multimedia Technology*, New York, NY, USA, Nov. 2022, pp. 84–89. doi: 10.1145/3551708.3556202.
[77] M. Khajah, R. Wing, R. Lindsey, and M. Mozer, “Incorporating Latent Factors Into Knowledge Tracing To Predict Individual Differences In Learning,” presented at the International Conference on Educational Data Mining, EDM 2014, 2014. Accessed: Nov. 08, 2021. [Online]. Available: https://www.semanticscholar.org/paper/Incorporating-Latent-Factors-Into-Knowledge-Tracing-Khajah-Wing/9af7938fd695fd2c90e566923828c6336d3c2292
[78] M. Khajah, Y. Huang, J. P. González-Brenes, M. C. Mozer, and P. Brusilovsky, “Integrating knowledge tracing and item response theory: A tale of two frameworks,” in *CEUR Workshop Proceedings*, Jan. 2014, vol. 1181, pp. 7–15. Accessed: Nov. 08, 2021. [Online]. Available: https://d-scholarship.pitt.edu/26044/
[79] Z. A. Pardos, Y. Bergner, D. T. Seaton, and D. E. Pritchard, “Adapting Bayesian Knowledge Tracing to a Massive Open Online Course in edX,” in *Proceedings of the 6th International Conference on Educational Data Mining, Memphis, Tennessee, USA, July 6-9, 2013*, Memphis, Tennessee, USA, 2013, pp. 137–144. Accessed: Dec. 03, 2021. [Online]. Available: http://www.educationaldatamining.org/EDM2013/papers/rn_paper_21.pdf
[80] S. P. Bhatt, J. Zhao, C. Thille, D. Zimmaro, and N. Gattani, “Evaluating Bayesian Knowledge Tracing for Estimating Learner Proficiency and Guiding Learner Behavior,” in *L@S’20: Seventh ACM Conference on Learning @ Scale, Virtual Event, USA, August 12-14, 2020*, Virtual , USA, 2020, pp. 357–360. doi: 10.1145/3386527.3406746.
[81] M. Eagle, A. T. Corbett, J. C. Stamper, and B. M. McLaren, “Predicting Individualized Learner Models Across Tutor Lessons,” in *Proceedings of the 11th International Conference on Educational Data Mining, EDM 2018, Buffalo, NY, USA, July 15-18, 2018*, Buffalo, NY, USA, 2018. Accessed: Dec. 03, 2021. [Online]. Available: http://educationaldatamining.org/files/conferences/EDM2018/papers/EDM2018_paper_217.pdf
[82] M. Yudelson, “Individualization of Bayesian Knowledge Tracing Through Elo-infusion,” in *Artificial Intelligence in Education - 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14-18, 2021, Proceedings, Part II*, Utrecht, The Netherlands, 2021, vol. 12749, pp. 412–416. doi: 10.1007/978-3-030-78270-2_73.
[83] J. Zhu, Y. Zang, H. Qiu, and T. Zhou, “Integrating Temporal Information Into Knowledge Tracing: A Temporal Difference Approach,” *IEEE Access*, vol. 6, pp. 27302–27312, 2018, doi: 10.1109/ACCESS.2018.2833874.
[84] Y. Xu, K. Chang, Y. Yuan, and J. Mostow, “EEG Helps Knowledge Tracing !,” 2014. Accessed: Dec. 03, 2021. [Online]. Available: https://www.semanticscholar.org/paper/EEG-Helps-Knowledge-Tracing-!-Xu-Chang/645cb28a5d3802c14305ae239a27dd3506ec71f9
[85] S. Corrigan, T. Barkley, and Z. A. Pardos, “Dynamic Approaches to Modeling Student Affect and its Changing Role in Learning and Performance,” in *User Modeling, Adaptation and Personalization - 23rd International Conference, UMAP 2015, Dublin, Ireland, June 29 - July 3, 2015. Proceedings*, Dublin, Ireland, 2015, vol. 9146, pp. 92–103. doi: 10.1007/978-3-319-20267-9_8.
[86] S. Spaulding, G. Gordon, and C. Breazeal, “Affect-Aware Student Models for Robot Tutors,” in *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems, Singapore, May 9-13, 2016*, 2016, pp. 864–872. Accessed: Dec. 03, 2021. [Online]. Available: http://dl.acm.org/citation.cfm?id=2937050
[87] M. A. Rau and Z. A. Pardos, “Adding eye-tracking AOI data to models of representation skills does not improve prediction accuracy,” in *Proceedings of the 9th International Conference on Educational Data Mining, EDM 2016, Raleigh, North Carolina, USA, June 29 - July 2, 2016*, 2016, pp. 622–623. Accessed: Aug. 18, 2021. [Online]. Available: http://www.educationaldatamining.org/EDM2016/proceedings/paper_165.pdf

[88] C. Lin, S. Shen, and M. Chi, “Incorporating Student Response Time and Tutor Instructional Interventions into Student Modeling,” in *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization, UMAP 2016, Halifax, NS, Canada, July 13 - 17, 2016*, 2016, pp. 157–161. doi: 10.1145/2930238.2930291.

[89] C. Lin and M. Chi, “Intervention-BKT: Incorporating Instructional Interventions into Bayesian Knowledge Tracing,” in *Intelligent Tutoring Systems*, Cham, 2016, pp. 208–218. doi: 10.1007/978-3-319-39583-8_20.

[90] J. I. Lee and E. Brunskill, “The Impact on Individualizing Student Models on Necessary Practice Opportunities,” in *Proceedings of the 5th International Conference on Educational Data Mining, Chania, Greece, June 19-21, 2012*, Chania, Greece, 2012, pp. 118–125. Accessed: Dec. 03, 2021. [Online]. Available: http://educationaldatamining.org/EDM2012/uploads/procs/Full_Papers/edm2012_full_11.pdf

[91] Z. A. Pardos, S. Trivedi, N. T. Heffernan, and G. N. Sárközy, “Clustered Knowledge Tracing,” in *Intelligent Tutoring Systems - 11th International Conference, ITS 2012, Chania, Crete, Greece, June 14-18, 2012. Proceedings*, Chania, Crete, Greece, 2012, vol. 7315, pp. 405–410. doi: 10.1007/978-3-642-30950-2_52.

[92] Y. Wang and N. T. Heffernan, “The Student Skill Model,” in *Intelligent Tutoring Systems - 11th International Conference, ITS 2012, Chania, Crete, Greece, June 14-18, 2012. Proceedings*, Chania, Crete, Greece, 2012, vol. 7315, pp. 399–404. doi: 10.1007/978-3-642-30950-2_51.

[93] G. Gorgun and O. Bulut, “Considering Disengaged Responses in Bayesian and Deep Knowledge Tracing,” in *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners’ and Doctoral Consortium*, Cham, 2022, pp. 591–594. doi: 10.1007/978-3-031-11647-6_122.

[94] J. E. Beck and J. Sison, “Using Knowledge Tracing to Measure Student Reading Proficiencies,” in *Intelligent Tutoring Systems, 7th International Conference, ITS 2004, Maceiò, Alagoas, Brazil, August 30 - September 3, 2004, Proceedings*, Maceiò, Alagoas, Brazil, 2004, vol. 3220, pp. 624–634. doi: 10.1007/978-3-540-30139-4_59.

[95] M. H. Falakmasir, M. Yudelson, S. Ritter, and K. R. Koedinger, “Spectral Bayesian Knowledge Tracing,” in *Proceedings of the 8th International Conference on Educational Data Mining, EDM 2015, Madrid, Spain, June 26-29, 2015*, Madrid, Spain, 2015, pp. 360–363. Accessed: Dec. 03, 2021. [Online]. Available: http://www.educationaldatamining.org/EDM2015/proceedings/short360-363.pdf

[96] J. E. Beck, “Difficulties in inferring student knowledge from observations (and why you should care).” 2007.

[97] J. E. Beck and K. Chang, “Identifiability: A Fundamental Problem of Student Modeling,” in *User Modeling 2007, 11th International Conference, UM 2007, Corfu, Greece, June 25-29, 2007, Proceedings*, Corfu, Greece, 2007, vol. 4511, pp. 137–146. doi: 10.1007/978-3-540-73078-1_17.

[98] S. Doroudi and E. Brunskill, “The Misidentified Identifiability Problem of Bayesian Knowledge Tracing,” International Educational Data Mining Society, Jun. 2017. Accessed: Mar. 01, 2023. [Online]. Available: https://eric.ed.gov/?id=ED596611

[99] PSLC DataShop, “Educational Data Mining Challenge - KDD Cup 2010,” 2010. https://pslcdatashop.web.cmu.edu/KDDCup/downloads.jsp (accessed Dec. 08, 2021).

[100] K. Chang, J. E. Beck, J. Mostow, and A. T. Corbett, “A Bayes Net Toolkit for Student Modeling in Intelligent Tutoring Systems. Intelligent Tutoring Systems,” in *Proceedings of the 8th International Conference on Intelligent Tutoring Systems, Jhongli*, 2006, pp. 104–113.

[101] K. P. Murphy, “The Bayes Net Toolbox for MATLAB,” *Computing Science and Statistics*, vol. 33, p. 2001, 2001.

[102] Y. Xu, M. J. Johnson, and Z. A. Pardos, “Scaling cognitive modeling to massive open environments,” in *In Proceedings of the Workshop on Machine Learning for Education at the 32nd International Conference on Machine Learn- ing (ICML)*, 2015.

ABBREVIATION LIST

AIS	Adaptive Instructional Systems
BKT	Bayesian Knowledge Tracing
CAI	Computer Assisted Instruction
EDM	Educational Data Mining
ILE	Intelligent Learning Environments
ITS	Intelligent Tutoring Systems
KC	Knowledge Component
LA	Learning Analytics
LMS	Learning Management Systems
ML	Machine Learning
MOOC	Massive Open Online Courses
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis

ABSTRACT

The quality of an Artificial Intelligence-based tutoring system is its ability to observe and interpret student behaviour to infer the preferences and needs of an individual student. The student model enables a comprehensive representation of student knowledge and affects the quality of the other Intelligent Tutoring System's (ITS) components. The Bayesian Knowledge Tracing (BKT) model is one of the first machine learning-based and widely investigated student models due to its interpretability and ability to infer student knowledge. The past 25 years have seen increasingly rapid advances in the field, so this systematic review deals with the BKT model enhancements by using the PRISMA guidelines and a unique set of criteria, including 13 aspects of enhancements and computational methods. Also, the study reveals two types of evaluation approaches found in the literature, including the prediction of student answers and the ability to estimate knowledge mastery. Overall, the most frequently investigated enhancements referred to the aspects added to the vanilla BKT model, including Student characteristics, Tutor interventions, and Domain knowledge properties, as well as the Question difficulty aspect pre-assumed as unique and constant. Regarding the used computational methods, the Expectation-Maximization algorithm practically became the standard in estimating BKT parameters. While the enhanced BKT models generally overperformed the vanilla model in predicting the student answer by using the measures such as RMSE (Root Mean Square Error), AUC-ROC (Area Under Curve, Receiver Operating Characteristics curve) and Accuracy, only a few studies further investigated the systems' estimations of knowledge mastery. The most frequently used educational platforms encompassed ITSs and Massive Open Online Courses (MOOCs), as well as simulated environments. The future work will focus on the adaptive and individualized formative tests based on BKT models with the aim to facilitate higher learning engagement and performance. In this sense, we seek for the student model that will enable the prediction of the overall student performance based on the weekly estimations of knowledge mastery.

SAŽETAK

Kvaliteta sustava za poučavanje koji se temelji na umjetnoj inteligenciji je njegova sposobnost promatranja i interpretacije učenikovog ponašanja s ciljem zaključivanja o individualnim potrebama. Model učenika, kao strukturna komponenta sustava za poučavanje, pruža cjelovit prikaz znanja i utječe na kvalitetu ostalih komponenti sustava. Bayesian Knowledge Tracing (BKT) model jedan je od prvih modela učenika koji se temelji na strojnom učenju i čest je predmet istraživanja zbog svoje interpretabilnosti i sposobnosti zaključivanja o znanju učenika. S obzirom da se model istražuje već više od 25 godina, ovdje se prezentira sustavan pregled poboljšanja osnovnog BKT modela. Pregled koristi Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) smjernice i jedinstveni skup kriterija koji se sastoji od 13 aspekata poboljšanja te korištenih računalnih metoda. Također, u pregledu se daje prikaz pristupa vrednovanju BKT modela, kao i uključene obrazovne platforme te statističke mjere. Rezultati pokazuju da poboljšani BKT modeli najčešće istražuju individualne karakteristike učenika, intervencije sustava, karakteristike područnog znanja, kao i uvođenje različite težine pitanja koja je u osnovnom modelu pretpostavljena kao jedinstvena. U kontekstu korištenih računalnih metoda, Expectation-Maximization metoda je praktički postala standard u procjeni parametara BKT modela. Vrednovanje BKT modela najčešće uključuje predviđanje učenikovog odgovora koristeći mjere poput RMSE (Root Mean Square Error), AUC-ROC (Area Under Curve, Receiver Operating Characteristics curve) i točnosti (Accuracy), dok se rjeđe istražuje sposobnost BKT modela pri procjeni ukupnog učenikovog znanja. Najčešće istraživane obrazovne platforme uključuju Inteligentne tutorske sustave (Intelligent Tutoring Systems, ITS) i Masovne otvorene online tečajeve (Massive Open Online Courses, MOOC), dok nekolicina istraživanja koristi i simulirane podatke. U daljnjem istraživanju ćemo se usredotočiti na adaptivne i individualizirane formativne testove temeljene na BKT modelima s ciljem povećanja angažmana u učenju te ukupne uspješnosti učenika. Istražit ćemo sposobnost takvog modela učenika temeljenog na tjednim komponentama znanja u predviđanju ukupnog znanja učenika.