User Modeling and User-Adapted Interaction
https://doi.org/10.1007/s11257-023-09389-4

Twenty-five years of Bayesian knowledge tracing: a systematic review

Ines Šarić-Grgić[1] · Ani Grubišić[1] · Angelina Gašpar[2]

Received: 30 March 2022 / Accepted in revised form: 25 November 2023

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 model (BKT) 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 Twenty-five 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 extended the vanilla BKT model by including student characteristics and tutor interventions. The educational context-based enhancements of domain knowledge properties, question difficulty and architectural prior knowledge were also frequently investigated enhancements. 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 by correlating it

✉ Ines Šarić-Grgić
ines.saric@pmfst.hr

Ani Grubišić
ani.grubisic@pmfst.hr

Angelina Gašpar
agaspar@kbf-st.hr

1 Faculty of Science, University of Split, Split, Croatia

2 Catholic Faculty of Theology, University of Split, Split, Croatia

Published online: 27 January 2024

to knowledge on post-tests. The most frequently used educational platforms included ITSs, Massive Open Online Courses (MOOCs) and simulated environments.

Keywords Student modelling · Bayesian knowledge tracing · Intelligent tutoring systems · Educational data mining

1 Introduction

In the traditional face-to-face learning environment, teachers use various methods (class discussions/observations, questioning, assignments, tests, and individual/group work) to assess students' learning progress and make decisions about their instructional strategies. 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*—a one-on-one tutor–student interaction in which technology mimics human teachers. Such an environment enables personalized and adaptive tutoring, considering potential difficulties and misconceptions.

The researchers of the interdisciplinary field of cognitive science, artificial intelligence, and educational technology have computerized teaching and learning by developing various types of *educational platforms* since the 1960s, e.g. computer assisted instruction (CAI) systems (Skinner 1954), *intelligent tutoring systems (ITS)* (Sleeman & Brown 1982), intelligent learning environments (ILE) (Wenger 1987), adaptive instructional systems (AIS) (DeFalco & Sinatra 2019), etc. Although different, all these platforms target adaptive and intelligent behaviour. The extensive research on ITS confirmed its standard architecture, consisting of the domain knowledge module (includes information about what to learn), the student module (what is the current knowledge of a student), the tutoring module (how to teach), and the communication module (tutor–student interface) (Freedman et al. 2000; Nkambou et al. 2010; Nwana 1990; Self 1974; Woolf 1992). The 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 other modules.

The main challenge of the student model is the *knowledge inference* process that aims to estimate the knowledge level during the tutoring process. In student–tutor interaction, the tutor maintains an estimate of the probability that the student has learned each of the rules in the ideal model, in the process called *knowledge tracing* (Corbett & Anderson 1995). Over the years, researchers used different machine Learning (ML) techniques to model student knowledge. The probabilistic *Bayesian Knowledge Tracing (BKT) is the most investigated approach,* first proposed by Corbett and Anderson (1995). The BKT uses the Hidden Markov model as a specific Bayesian network that models the student's knowledge mastery as a latent variable. In the BKT model, knowledge components (KC) represent skills or concepts the student learns.

For each KC, the model maintains a set of parameters to calculate the overall probability of knowledge mastery. Since the vanilla term refers to technology not customized or updated from its standard form, the BKT model proposed by Corbett and Anderson is later called the *vanilla BKT model.*

Researchers have been able to validate and refine the BKT model for over Twenty-five Years due to the availability of ITS's and the wealth of data they provide. They proposed various BKT enhancements, but several studies focused on student modelling approaches (Anouar Tadlaoui et al. 2016; Brusilovsky & Millán, 2007; Chrysafiadi & Virvou 2013; Desmarais & Baker 2012; Harrison & Roberts 2012; Kurup et al. 2016; Liu et al. 2021b; Pavlik et al. 2013; Pelánek 2017; Ramírez Luelmo et al. 2021; Sani et al. 2016; Vandewaetere et al. 2011; Zafar & Ahmad 2013). The educational technology and data analytics literature still lacks a comprehensive and organized compilation of all existing research on the family of BKT models, including the latest advancements. This review study applies the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (Moher et al. 2009), including (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.

The 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 data collected from educational platforms and performance measures were part of the research on the BKT enhancements?

This study can contribute to a more detailed understanding of how to advance BKT models. It can be appealing to researchers and practitioners in this interdisciplinary field, who look for systematic reviews and potential directions for new models. Since the researchers typically compared the enhanced BKT models to the vanilla model, the unique set of vanilla assumptions provided a basis for the review criteria.

The remainder of this paper is structured as follows: Sect. 2 deals with the theoretical background and Sect. 3 describes the vanilla BKT model. Section 4 provides the methodology, including the research questions and the proposed criteria of the review process. The results and discussion are presented in Sects. 5 and 6, followed by the conclusion in Sect. 7.

2 Theoretical background

The background section aims to present (i) the use of BKT in the previous review studies on student modelling, (ii) the overview of ML techniques used for student modelling and (iii) the comparison of the Bayesian network-based BKT model to the logistic regression-based and neural network-based models.

This summary of previous literature on student modelling refers to *overviews* (Brusilovsky & Millán, 2007; Desmarais & Baker 2012; Harrison & Roberts 2012; Kurup et al. 2016; Vandewaetere et al. 2011)**,** *reviews* (Anouar Tadlaoui et al. 2016; Liu et al. 2021b; Pavlik et al. 2013; Pelánek 2017; Sani et al. 2016; Zafar & Ahmad 2013) and *systematic literature reviews* (Chrysafiadi & Virvou 2013; Ramírez Luelmo et al. 2021). Table 1 provides the complete list of proposed taxonomies of student modelling approaches in descending chronological order, including references, a research focus and the proposed taxonomy of student modelling approaches. The BKT-related student modelling approaches are in italics.

Table 1 Proposed taxonomies of student modelling approaches

References	Research focus	Proposed taxonomy of student modelling approaches
Liu et al. (2021b)	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 Adaptive Educational Systems	Overlay; Stereotype; Machine Learning; Plan Recognition; Differential; Perturbation; *Bayesian Networks*
Sani et al. (2016)	The review focused on 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 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 Intelligent Tutoring Systems	Overlay models (Rule Space models, *Model Tracing models,* Constraint-based models); Knowledge Space models; Dialogue models; Programmed Branching; State and Trait

Table 1 (continued)

References	Research focus	Proposed taxonomy of student modelling approaches
Chrysafiadi & Virvou (2013)	The systematic review focused on 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 Adaptive Hypermedia and Adaptive Educational Systems	Overlay; *Uncertainty-based modelling*

We emphasize that the previous research studies on student modelling approaches originated from the research community of ITS and adaptive instruction. We acknowledge that we are aware of a large body of work on cognitive diagnostic modelling and psychometrics that come from different research communities (e.g. Item Response Theory, DINA and its family of models), but we do not additionally include these approaches in the review of ML methods for student modelling.

Each study in Table 1 gave an overview of student modelling approaches related to specific educational platforms, adaptive behaviour or used techniques. In the early reviews, ML techniques were already the basis for different student modelling approaches. Over time, new techniques complemented the previous student modelling taxonomies. 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 enabling machines to learn through experience and using data (Mitchell 1997). ML techniques used for student modelling enhance the adaptiveness and intelligence of educational platforms. Those identified in the already mentioned research studies include Bayesian networks, logistic regression, neural networks, support vector machines, fuzzy logic, and matrix factorization (Table 2).

Moreover, Liu et al. (2021b) classified the student modelling approaches as probabilistic, logistic, and deep learning-based models. *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. *Logistic* models, such as learning factor analysis (Cen et al. 2006)

Table 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. (2021b)	+	+	+	–	–	–
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)	+	–	–	–	+	–

and performance factor analysis (Pavlik et al. 2009), predict the probability of student performance by learning function, typically a logistic function. The last group of knowledge tracing approaches are *deep learning* models, based on neural networks (Piech et al. 2015). We support the shortest high-level classification of student modelling approaches, but there is inconsistency with this classification, since the logistic models can be considered probabilistic.

Ramirez Luelmo et al. (2021) investigated ML techniques employed in student modelling from 2015 to 2020. Their research results indicate the most common ML techniques as BKT (18 applications), deep knowledge tracing (13 app.), long-short term neural networks (12 app.), Bayesian networks (11 app.), support vector machines (7 app.), dynamic key-value memory networks (7 app.), and performance factor analysis (6 app.).

Overall, Bayesian networks are the continuously investigated ML technique used for student modelling. The vanilla BKT based on the hidden Markov model is the most representative and unique student modelling approach, considered by researchers in the field as a baseline.

As for the preference between the probabilistic Bayesian network-based and logistic regression-based models, researchers often prefer one model but provide no rationale behind their choices (Pelánek 2017). 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 (Montero et al. 2018). 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 (Badrinath et al. 2021; Khajah et al. 2016). Based on Bayesian statistics, the BKT model assumes knowledge and performance nodes with binary states and is more interpretable than the neural network-based model. 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 in 2022,
- Reviews the BKT enhancements and evaluation approaches, including datasets from educational platforms and the performance measures found in the literature.

3 Bayesian knowledge tracing

This section provides a brief description of the vanilla BKT model. It is one of the first ML knowledge tracing models introduced by Corbett and Anderson (1995) and is considered the first significant milestone of the Educational Data Mining research field (Baker & Inventado 2014).

The vanilla model results from the work of the ACT Programming Tutor and reflects the ACT-R cognitive theory (Adaptive Control of Thought–Rational) (Anderson 1993), which states that mastering a complex skill implies its components and the use of the Bayesian computational procedure identified in Atkinson's work (Atkinson 1972).

A Bayesian network is a probabilistic graphical model for representing knowledge about an uncertain domain, where each node represents a random variable and directed edges between nodes indicate probabilistic dependencies between these variables. A Hidden Markov Model (HMM) is a special Bayesian Network for tracing not directly observable (hidden) nodes using the observable node states. In BKT, hidden nodes represent *student knowledge* and observable nodes represent *student performance*. Both types of nodes are assumed to be binary, including the *unlearned* and *learned* knowledge states and the *correct* and *incorrect* performance states.

Figure 1 shows the hidden student $knowledge_t$, $t \in \{1, 2, \ldots T\}$ and observable *student* $performance_t$ nodes, $t \in \{1, 2, \ldots T\}$ of the vanilla HMM. While $P(L_0)$ is the initial probability of knowledge before any opportunity of applying it (prior knowledge), there are also *transition and emission probabilities.* The transition probabilities refer to the probability $P(T)$ of a knowledge transitioning from *unlearned* state to *learned* state 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 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.

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 $P(L_t)$, $t \geq 0$.

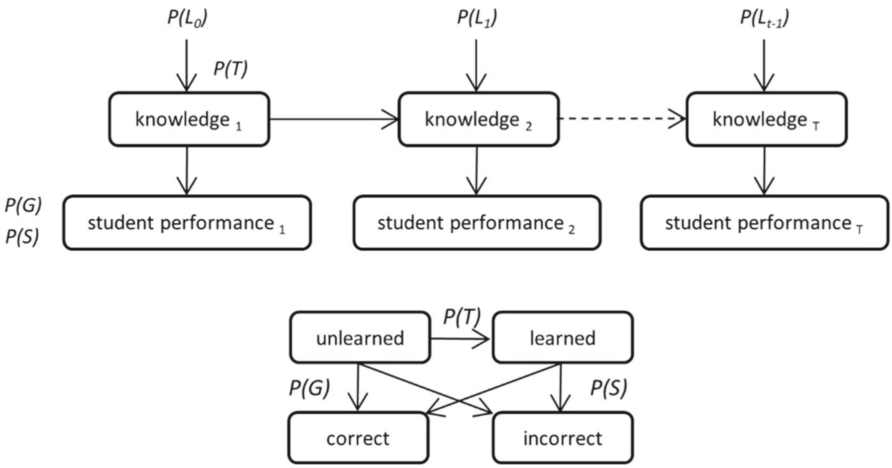

Fig. 1 The vanilla BKT model and its instantiation process—based on Zhang and Yao (2018)

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- $P(G)$

Fig. 2 BKT parameters in a matrix form (Yudelson et al. 2013)

The model updates the probability $P(L_t)$ after each opportunity to apply knowledge given an observed correct or incorrect response as follows:

$$P\left(L_{t-1}|\text{Correct}_t\right) = \frac{P(L_{t-1})(1 - P(S))}{P(L_{t-1})(1 - P(S)) + (1 - P(L_{t-1}))P(G)} \tag{1}$$

$$P\left(L_{t-1}|\text{Incorrect}_t\right) = \frac{P(L_{t-1})P(S)}{P(L_{t-1})P(S) + (1 - P(L_{t-1}))(1 - P(G))} \tag{2}$$

If $evidence_t \in \{\text{Correct}_t, \text{Incorrect}_t\}$ represents the observable correctness of a student's answer after an opportunity t to apply knowledge, the updated probability for the following time step is defined as:

$$P(L_t) = P\left(L_{t-1}|\text{evidence}_t\right) + (1 - P(L_{t-1}|\text{evidence}_t)) * P(T) \tag{3}$$

Using the evidence from the current step, the model first calculates the probability that the student knew the answer before making an attempt. Then, taking this into account, it computes the likelihood that the student learned it after making the attempt. The previous Eqs. (1), (2) and (3) are based on the original BKT publication (Corbett & Anderson 1995), and the similar notation appears in the work by Zhang and Yao (2018).

Regarding the BKT parameter estimation procedure, Corbett and Anderson (1995) discussed individualization per skill and individualization per student of all four BKT parameters. The individualized BKT model resulted in a better correlation between actual and expected accuracy across student results than the non-individualized BKT model whose accuracy of predicting student test scores (after a period of working with a tutoring system) did not improve tangibly (Yudelson et al. 2013). 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 in this work is in line with the *PRISMA guidelines* (Moher et al., 2009) 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 on the vanilla BKT model enhancements published in the relevant scientific databases until 2022. The implementations of the BKT enhancements could proceed in two directions, including the Bayesian network architecture/educational context and new computational methods.

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 (the final refinements made in March 2023). The search strategy included the expression *knowledge tracing* and versions of the words *Bayes* and *probabilistic*, contained in the publication abstracts. Due to the extensiveness of the Google Scholar database, we searched the publication titles using the expression *Bayesian knowledge tracing*. Table 3 shows the search details with the number of publications. Figure 3 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 84 excluded publications, which were off-topic, written in languages other than English or as a programming code.

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

The full-text reading phase included the remaining 56 publications in which we found 17 references and cited in this review. Those publications were part of specific events (e.g. Chang et al. (2006b) presented at 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. (2008b)).

Finally, this systematic review includes 73 publications and the original BKT publication.

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, in which 'Other' denotes sources that contributed to the review study with a single publication. The most common sources were scientific conferences, including the International Conference on Educational Data Mining (EDM), International Conference of Intelligent Tutoring Systems (ITS), International Conference of Artificial Intelligence in Education (AIED), International Conference on User Modelling, Adaptation, and Personalization (UMAP), ACM Conference on Learning at Scale (L@S), International Conference on Learning Analytics & Knowledge (LA&K), IEEE Conference on Big Data (IEEE BigData) and the User Modelling and User-Adapted Interaction (UMUAI)

Table 3 Database search details

Database	Search query (—2022)	#
Web of science	"knowledge tracing" (Abstract) AND (bayes* OR probab*) (Abstract)	89
Scopus	ABS ("knowledge tracing") AND ABS ((bayes* OR probab*))	200
ACM	[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

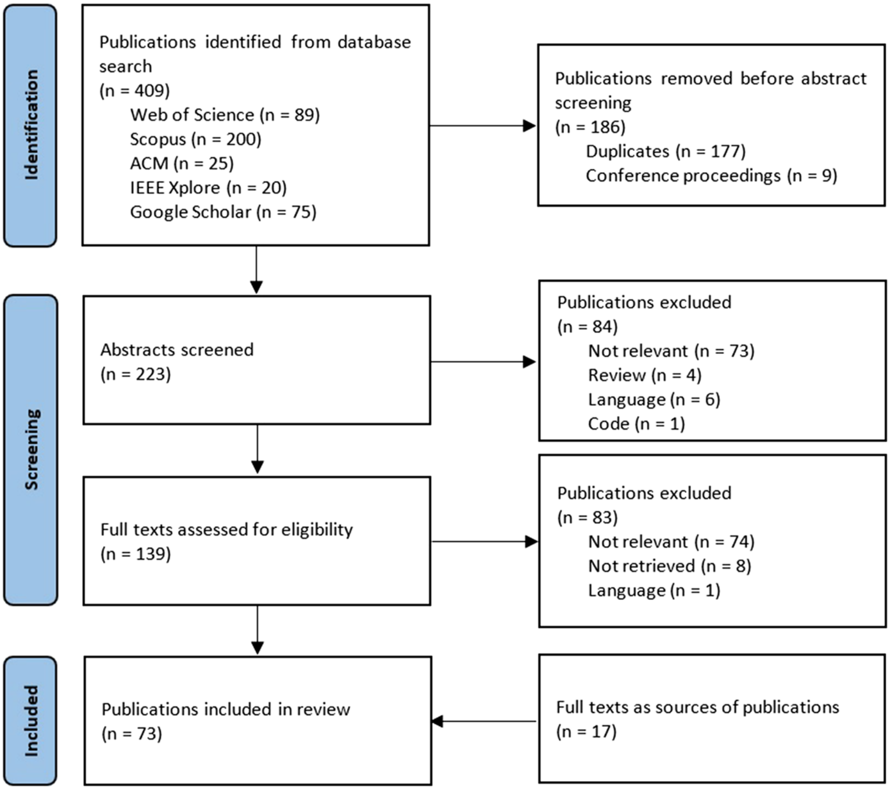

Fig. 3 PRISMA flow diagram of the publication identification and screening process

Table 4 Heatmap of the most frequent sources of publications in the research of BKT enhancements

Conf./Journal	19 95	20 04	20 06	20 07	20 08	20 09	20 10	20 11	20 12	20 13	20 14	20 15	20 16	20 17	20 18	20 20	20 21	20 22	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**

Journal. There was an increase in publications in 2008, 2010, and between 2013 and 2018.

To address *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 1995. 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 component 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 refer to 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) refer to 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 RQ1, we reviewed the computational methods used in the enhanced BKT models.

Regarding *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 the reported results, this review study provides more insights into the evaluation approaches.

5 Results

This systematic overview of BKT enhancement aspects encompassed the identified research studies, which resulted in 62 enhanced BKT models. We noted some publications as multiple sources of the single enhanced BKT model (e.g. Beck et al. 2008; Chang et al. 2006b). For more than one source publication per model, we considered the year of the earlier publication as a model source year. Appendix A provides the complete list of enhanced BKT models, source publications, model name, and a short description.

While some of the BKT models addressed the architectural and educational context-based properties of the vanilla BKT model (A01-A10), some enhancements extended the characteristics of the vanilla BKT model. Both types of enhancement aspects could also propose new computational methods. Therefore, we found it important to analyse architectural and educational context-based enhancements and computational methods separately.

Section 5.1 reviews the enhancements of the architectural and educational context-based properties, including the specific enhancement aspect of each of the four BKT parameters (A03-A06). Section 5.2 overviews computational methods generally used for the parameter estimation, followed by the evaluation approaches of enhanced BKT models in Sect. 5.3.

5.1 The architectural and context-based enhancements (RQ1)

To review the enhanced BKT models, we proposed the enhancement criteria in line with the vanilla BKT model assumptions. The enhancement criteria resulted from an iterative analysis of the identified research studies and represented a unique way of classifying the BKT enhancements. Besides those criteria found in the vanilla BKT model (enhancement aspects EA01-EA10 in Table 5), some enhancements extended the vanilla BKT model with new aspects. Additional vanilla BKT enhancement aspects included student characteristics (EA11), tutor interventions (EA12), and noise in data (EA13). Table 5 shows the complete list of BKT enhancements and the related vanilla model assumptions.

Although each change in the Bayesian network architecture directly implied the update of BKT parameters, EA03-EA06 criteria encompassed BKT models focusing on the prior knowledge, guessing, slipping and learning BKT parameters, e.g. Contextual Guess and Slip method (Baker et al. 2008a). A yearly heatmap (Table 6)

Table 5 Enhancement criteria used to review BKT models

BKT Enhancement Aspects (EA)		Vanilla BKT model assumptions (A)	
EA01	Knowledge states	A01	Knowledge component node in the Bayesian network includes the binary *learned* and *unlearned* state
EA02	Performance states	A02	Performance node in the Bayesian network includes the binary *correct* and *incorrect* state
EA03	Prior knowledge	A03	The prior knowledge probability is defined by experts and per skill
EA04	Guessing	A04	There is a probability of guessing defined by experts per skill
EA05	Slipping	A05	There is a probability of slipping defined by experts per skill
EA06	Learning	A06	The learning transition probability is defined by experts per skill
EA07	Forgetting paradigm	A07	The *no-forgetting paradigm* is followed meaning that there is no transition from learned to unlearned state
EA08	Domain knowledge properties	A08	Domain knowledge fractionates into *independent* knowledge components
EA09	Question difficulty	A09	Questions of each knowledge component are of *equal difficulty*
EA10	Multiple attempts	A10	The *first attempt* to answer the question counts during the modelling process
EA11	Student characteristics	Not included	
EA12	Tutor interventions	Not included	
EA13	Noise in data	Not included	

Table 6 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
EA01	Knowledge states			1								1	2			1		5
EA02	Performance states			1	1					1	2					1		6
EA03	Prior knowledge				1			2	1	2	1	1						8
EA04	Guessing			1	1	1						1	1					5
EA05	Slipping			1	1	2						1	1					6
EA06	Learning							3					1					4
EA07	Forgetting paradigm		1	1		1				1	1		1					6
EA08	Domain know. prop.							1	3	1	3	1		1			2	12
EA09	Question difficulty					1		1	3	1	1	1	1					9
EA10	Multiple attempts			1	1			1	1	1	1				1			7
EA11	Student characteristics			1	1		3		4	1	4		4			1	1	20
EA12	Tutor interventions		1	2	2					1	2	1	1					10
EA13	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**

presents the review of BKT enhancements. 54 enhanced BKT models addressed 101 architectural and educational context-based properties of the vanilla BKT model.

The first enhanced BKT model emerged in 2004, a decade after the vanilla model. There was decrease in the research after 2018, probably due to the COVID-19 pandemic. Much research work focused on enhancements to the vanilla model as student characteristics (20 research studies), domain knowledge properties (12 r.s.), tutor interventions (10 r.s.), and question difficulty (9 r.s.). The most investigated enhancements extended the vanilla BKT model, more specifically student characteristics and tutor interventions. The most investigated educational context-based enhancements referred to domain knowledge properties and question difficulty, while the most frequently investigated architectural enhancement included the prior knowledge.

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 7, and variations found in a single research study are summarized as 'Other'.

Table 7 The variations of investigated BKT enhancements

#	BKT enhancement aspects (EA)	# Enhanced BKT models
1	EA11 Student characteristics	9
2	EA08 Domain knowledge properties	7
3	EA03 Prior knowledge	5
4	EA04 Guessing, EA05 Slipping, EA11 Student characteristics, EA12 Tutor interventions	3
5	EA09 Question difficulty, EA11 Student characteristics	2
6	EA11 Student characteristics, EA12 Tutor interventions	2
7	EA13 Noise in data	2
8–31	Other	24

Among 31 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.

Table 8 shows the research related to each BKT enhancement and the discussion section elaborates on the approaches.

Table 8 Enhanced BKT models per proposed criteria

BKT Enhancement Aspects (EA)		BKT models
EA01	Knowledge states	Halpern et al. (2018), F. Liu et al. (2021a), Schodde et al. (2017), Yudelson et al. (2008), K. Zhang & Yao (2018)
EA02	Performance states	David et al. (2016), F. Liu et al. (2021a), Ostrow et al. (2015), Y. Wang et al. (2010), Y. Wang & Heffernan (2013), Z. Wang et al. (2016), Yudelson et al. (2008)
EA03	Prior knowledge	Eagle et al. (2016a, b, Eagle et al. (2017), Nedungadi & Remya (2014, 2015), Pardos & Heffernan (2010), Song et al. (2015), S. Wang et al. (2017), Xu & Mostow (2013), Yudelson et al. (2013)
EA04	Guessing	Agarwal et al. (2018), Baker et al. (2008a, 2008b, 2010), Pardos & Heffernan (2011), Zhou et al. (2017)
EA05	Slipping	Agarwal et al. (2018), Baker et al. (2008a, 2008b, 2010), Pardos & Heffernan (2011), Qiu et al. (2011), Zhou et al. (2017)
EA06	Learning	Adjei et al. (2013), Baker et al. (2018), Sao Pedro et al. (2013), Yudelson et al. (2013)
EA07	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)
EA08	Domain know. prop	Chan et al. (2022), González-Brenes et al. (2014; Hawkins & Heffernan (2014), Huang et al. (2016), Huang & Brusilovsky (2016), Khajah et al. (2016), MacHardy (2015) MacHardy & Pardos (2015), Meng et al. (2019), Sao Pedro et al. (2013, 2014), Sun et al. (2022), Z. Wang et al. (2016, 2016)
EA09	Question difficulty	Baker et al. (2018), David et al. (2016), González-Brenes et al. (2014), 2014a, b; Ostrow et al. (2015), Pardos et al. (2013), Pardos & Heffernan (2011), Zhou et al. 2017)
EA10	Multiple attempts	Bhatt et al. (2020), González-Brenes et al. (2014), Pardos et al. (2013), Yudelson et al. (2008)

Table 8 (continued)

BKT Enhancement Aspects (EA)		BKT models
EA11	Student characteristics	Agarwal et al. (2018), Baker et al. (2008a, 2008b,2010), Corrigan et al. (2015), Eagle et al. (2018), Gorgun & Bulut (2022), Halpern et al. (2018), 2014a, b, Khajah et al. (2016), Lee & Brunskill (2012), Lin et al. (2016), Lin & Chi (2016), Nedungadi & Remya (2014), Pardos et al. (2012), Rau & Pardos (2016), Spaulding et al. (2016), Wang & Heffernan (2012), Xu et al. (2014), Yudelson (2021), Zhu et al. (2018)
EA12	Tutor interventions	Agarwal et al. (2018), Baker et al. (2008a, 2010,2008b), 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), Wang et al. 2010), Wang & Heffernan (2013), Yudelson et al. 2008)
EA13	Noise in data	Beck & Sison (2004), Falakmasir et al. (2015), Gorgun & Bulut (2022)

5.2 Computational methods (RQ1)

Computational methods used in the proposed BKT approaches generally referred to the estimation of BKT parameters. Some models did not have to interfere with the vanilla model assumptions, but primarily addressed the computational challenges, e.g. Dirichlet priors method (Beck 2007; Beck & Chang 2007). Overall, 56 enhanced BKT models reported the use of computational methods.

Table 9 presents the results of the review of computational methods used in the research of BKT enhancements with over 2 applications.

Computational methods enhanced the skill-based estimations of BKT parameters used in the vanilla model. The expectation–maximization method, first used in 2006,

Table 9 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

Table 10 The datasets collected from educational platforms and used in the research of BKT enhancements

Educational platforms	2004	2006	2007	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
JavaGuide									1	1		1							3
Reading Tutor		1	1							1									3
Andes Tutor										2									2
Inq-ITS									1	1									2
Robot Tutor												1	1						2
Other	1			1						2	1	6	2	3	1	1			18
Total (T)	1	1	1	2	1	4	4	3	8	12	7	14	4	8	3	1	3	3	**80**

Table 11 Performance measures used in the research of BKT enhancements

#	Performance measures with over 2 applications	# Enhanced BKT models
1	RMSE	28
2	AUC–ROC	23
3	Accuracy	20
4	MAE	8
5	Correlation	3
6–14	Other	13

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.).

5.3 Evaluation approaches (RQ2)

Concerning the evaluation approaches, we reviewed educational platforms and performance measures used in the research of BKT enhancements.

Table 10 shows a yearly heatmap of the educational platforms used in the reviewed publications. We summarized those platforms with a single application as 'Other' denotes platforms with a single application).

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

As for the domain, the examined datasets related to Math (38 r.s.), Language learning and Programming (per 7 r.s.), Genetics, Physics and Engineering (per 3 r.s.), Science (per 2 r.s.), and Medicine and Chemistry (per single r.s.).

Regarding the used performance measures, Table 11 shows the most frequently used performance measures in the research of BKT enhancements with over 2 applications.

The most frequently used performance measures included the RMSE measure (root mean square error, 28r.s.), the AUC–ROC (area under curve, receiver operating characteristics curve, 23 r.s.) and the accuracy measure (20 r.s.). These performance measures are frequently used metrics for classification tasks in the machine learning field.

6 Discussion

The following subsections discuss the BKT enhancements based on the proposed aspects (EA01-EA13) and computational methods (RQ2), as well as the evaluation approaches of enhanced models (RQ2).

6.1 BKT enhancements (RQ1)

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

6.1.1 Knowledge states (EA01)

While the vanilla BKT model used the binary states of the knowledge node, several approaches proposed additional states (Halpern et al. 2018; Liu et al. 2021a; Schodde et al. 2017; Yudelson et al. 2008; Zhang & Yao 2018).

Yudelson et al. (2008) first researched partially mastered knowledge as an additional state of the hidden knowledge node. Along with a hints as an additional performance state, forgetting, the aggregation of skills, and all attempts, the research results supported the extension of the knowledge states. From a machine learning perspective, additional knowledge states improve the BKT model likelihood in reaching a better agreement allowing for a different knowledge state at every time step.

To better quantify the system's uncertainty about the learner's skill in adaptive language tutoring, Schodde et al. (2017) investigated six states of the knowledge node based on the discrete values of mastery percentage belief (0%, 20%, 40%, 60%, 80%, 100%). Compared to the randomized training, the adaptive model proved more successful. However, there was no significant difference across experimental conditions in the case of the post-test results.

Zhang and Yao (2018) proposed trisection of the learning process using three-way decision theory. Besides the unlearned and learned knowledge states, they introduced the learning knowledge state. The results of the comparative experiments demonstrated the improved prediction accuracies over the vanilla model.

While Halpern et al. (2018) investigated the integration of the neuroimaging recordings (fMRI) into the knowledge tracing model and proposed the third permanently learned knowledge state. A more elaborated model of memory leveraged more subtle dynamics of the fMRI data.

Liu et al. (2021a) used fuzzy methods to capture fine-grained changes in student cognition; the type-1 method for fuzzification of concepts and relations and type-2 method for relations. The experimental results revealed the better performance results than the baseline bisection BKT models.

6.1.2 Performance states (EA02)

As for hidden knowledge node, the binary states of performance node were additionally investigated in the literature (David et al. 2016; Liu et al. 2021a; Ostrow et al. 2015; Wang et al. 2010; Wang & Heffernan 2013; Wang et al. 2016; Yudelson et al. 2008).

Besides the correct and incorrect state of the performance node, Yudelson et al. (2008) proposed the third state based on the use of hints during the learning process.

Along with hints, Wang et al. (2010) used the number of attempts to assign the partial credit score of student's answer. The results revealed that the partial credit improved the model fitting by only a small absolute value compared to the approach that considers mean previous binary performance. A high relative value was reported compared to the baseline approach based on the binary performance and student id.

By keeping the vanilla model architecture, Ostrow et al. (2015) investigated the partial credit using five credit bins (0, 0.03, 0.6, 0.7, 0.8, 1.0) based on the first response type (attempt, hint request or scaffold request), attempt count, and hint count. The model proposing partial credit and various problem difficulty achieved mixed performance results. While student level analysis of partial credit tabling method achieved better results, the analysis per problem performed about as well as the compared approaches.

The partial credit approach introduced by David et al. (2016) treated the evidence (the student's score for a question) as a weighting factor determining the extent to which the student's response was correct. It differed from the approach by Wang et al. (2010) since it did not rely on the expectation–maximization algorithm and Ostrow et al. (2015) approach, not using predefined methods to determine the posterior distributions. The model combining partial credit, item difficulty and multiple attempts achieved better results than compared models.

Wang et al. (2016) investigated the relations between subskills concerning the hierarchical properties (the structure) and temporal properties in the MOOC environment. Its enhanced architecture uses binary values (correct or incorrect) for the performance node and the third value (no observation) if there is no student answer. Based on the experimental results, the proposed models overperformed the vanilla model.

The fuzzy approach proposed by Liu et al. (2021a) addressed continuous score scenarios to extend the bisection representation of student performance.

6.1.3 Prior knowledge (EA03)

While Corbett and Anderson (1995) initially discussed fitting BKT parameters per skill or student, the prior knowledge enhancement aspect included additional approaches regarding individualized prior knowledge parameter. Several works investigated how the expert-based probability estimations used in the vanilla model could be further enhanced (Eagle et al. (2016a, b; Eagle et al. 2017; Nedungadi & Remya 2014, 2015;

Pardos & Heffernan 2010; Song et al. 2015; S. Wang et al. 2017; Xu & Mostow 2013; Yudelson et al. 2013).

To individualize the vanilla model, Pardos and Heffernan (2010) compared three approaches of the prior knowledge parameter, including the random values, cold start heuristic (the first response), and per cent correct heuristic approach. The proposed models fit better than the vanilla model on a significant fraction of the examined problem sets, and the per cent correct performed the best. Also, the best strategy used information from multiple skills to inform each student's prior.

Xu and Mostow (2013) proposed a higher order Item Response Theory model. The model approximated initial knowledge as students' one-dimensional (or low-dimensional) overall proficiency and combined it with estimated difficulty and discrimination of each skill to estimate the probability of knowing a skill before practising it. The proposed model overperformed those using parameter estimations per skill or per student.

Yudelson (2013) investigated how the inclusion of the individualized prior knowledge and the transitional learning BKT parameter would affect the model performance. The individualized models could significantly improve the accuracy of predicting the student success. An interesting finding was that adding student-specific probability of learning parameter proved more beneficial for the model accuracy than adding student-specific probability of prior knowledge.

Nedungadi and Remya (2014) proposed the enhanced BKT model based on individual priors for each student and skill and dynamic clustering of students based on changing learning ability. They used a clustering of students and skills based on a student and skill capability matrix to learn the prior skills and to deal with the cold start problem (the prediction for either new skills or new students). The proposed model achieved better experimental results than the compared models. The model was further enhanced with forgetting (Nedungadi & Remya 2015), which resulted in the additional and significant increase in the accuracy of the student prediction.

Song et al. (2015) investigated the prior knowledge parameter in the case of the first-encounter knowledge prediction. The proposed model relied on the user-based collaborative filtering and the reported experimental results proved acceptable.

Eagle et al. (2016a, b; Eagle et al. 2017) examined prior knowledge to predict performance in the first two practice opportunities. The proposed model relied on the performance in activities that naturally precede tutor modules: reading online instructional text and taking a conceptual knowledge pre-test. The study reported that the proposed model performed wellconcerning model fitting and learning efficiency.

In the model proposed by Wang et al. (2017), the Item Response Theory (IRT) model was used to estimate the students' initial knowledge status and then joined with the discrimination and difficulty of each skill to evaluate the probability of knowing a skill before training it. The model combining two theories performed better than the single models.

6.1.4 Guessing (EA04) and slipping (EA05)

Several research works focused on the guessing and slipping BKT parameters (Agarwal et al. 2018; Baker et al. 2008a, 2010, 2008b; Pardos & Heffernan 2011; Zhou et al.

2017). Only one research on BKT model investigated slipping without investigating the guessing parameter (Qiu et al. 2011).

Baker and colleagues (2008a, 2008b) proposed the Contextual Guess and Slip method (CGS-BKT) based on the dynamic estimation of parameters using the contextual information of the action–learning speed and history of help-seeking. In the later research (Baker et al. 2010), authors combined CGS-BKT method with the Brute-force method for the parameter estimation. Also, Agarwal et al. (2018) proposed iterative technique based on contextual guess and slip estimation that converges to stable estimates for skill-level guess and slip parameters.

Addressing the item difficulty aspect, Pardos and Heffernan (2011) proposed individualized guess and slip parameters per each question.

The model proposed by Zhou et al. (2017) assumed that the guessing and slipping probabilities on each step rely on the student ability of each subskill—the more subskills a step needs, the less possible a student could guess the right answer.

The only model that investigated slipping without guessing was found in the research by Qiu et al. (2011). The model focused on poor student performance on the next day of learning and, in that sense, assumed it as forgetting or slipping.

6.1.5 Learning (EA06)

Four studies on enhanced BKT models further researched the learning parameter of the network architecture (Adjei et al. 2013; Baker et al. 2018; Sao Pedro et al. 2013; Yudelson et al. 2013).

Adjei et al. (2013) examined if students learned the same amount of knowledge according to whether students get items correct or wrong. The proposed model assumed different learning rates based on the correctness of a student answer. Allowing learning rates from previous incorrect performances to be higher seems intuitive, but the experiments showed that this approach did not lead to better predictions.

Already mentioned research by Yudelson et al. (2013) specifically investigated the learning transition of the BKT's network architecture by incrementally allowing the skill learning probabilities. The model did not require changes in the Bayesian network and proved useful along with a wide range of existing gradient descent algorithms. Adding student-specific learning probability was more beneficial for the model accuracy than adding student-specific initial mastery probability.

Sao Pedro et al. (2013) proposed different learning rates to investigate the scaffolding provided by the system and the transfer of skills between divergent topics.

The model proposed by Baker et al. (2018) assumed that learning of content occurs periodically during the learning and assessment process, typically when video or other instructional materials are reached. Therefore, the moment for the application of the regular learning formula was updated. Combined with different guess and slip parameters for each skill, the model led to better results.

6.1.6 Forgetting paradigm (EA07)

While the vanilla model omitted the possibility of forgetting the learned knowledge, several works further investigated this assumption (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).

While investigating the measures of the effectiveness of tutor interventions, Chang et al. (2006b) enabled the forgetting parameter in the Bayesian network architecture. Even though the additional information from the tutor help did not yield a superior model fit, the reported results emphasized its informativeness.

Yudelson et al. (2008) attempted to learn the memory decay directly from the individual students' data using a logarithmic function. The observed time variable added to the network architecture was discretized logarithmic interval representing learning within and between problems.

Qiu et al. (2011) investigated the student responses where a day or more had elapsed since the previous response, assuming it was a forgetting or slipping parameter. The result showed a significant improvement in the overall prediction of the forgetting model, suggesting that forgetting is more likely a cognitive explanation for the data.

The models proposed by Nedungadi and Remya (2015) incorporated forgetting by updating the knowledge level for a skill based on the duration of the previous attempt at using the skill. The exponential decay function used the time lag between the prior skill to update the knowledge level. The results showed a significant increase in prediction accuracy over the compared models.

Based on the previous forgetting research that focuses on the next attempt at learning (Qiu et al. 2011, Khajah et al. (2016) focused on the forgetting that can occur on a much shorter time scale. The proposed model achieved positive results, even as good as the deep knowledge tracing model.

Halpern et al. (2018) investigated fMRI measurements and improved a learner's test performance prediction at a 72-h delay. The three-state model included forgetting as a decay transition between the learned and unlearned state of the knowledge node. There is no possibility of forgetting at a permanently learned knowledge state.

6.1.7 Domain knowledge properties (EA08)

The BKT enhancement aspect of domain knowledge properties encompassed the research on the relations between skills (Chan et al. 2022; 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; Sun et al. 2022; Z. Wang et al. 2016), relations between questions (Hawkins & Heffernan 2014), and the use of video content in MOOCs (MacHardy 2015; MacHardy & Pardos 2015; Wang et al. 2016).

Sao Pedro et al. (2013) investigated skill transfer between different topics. Learning rates for each topic allowed the skill transfer across topics. In the later research (2014), authors improved the methodology and confirmed the transfer of inquiry skills across examined topics.

Gonzales-Brenes et al. (2014) proposed the Feature Aware Student knowledge Tracing (FAST) model that allows subskills. By enabling the binary subskill indicator and using logistic regression, the model estimates slip and guess probabilities as a function of the subskills in the practice opportunity. The research suggested that the subskills modelling is important and reported improved results over the compared approaches.

Huang and coauthors (2016) examined the importance of multiple skill combinations to gain extra knowledge or produce difficulty not captured by contributing individual skills. The model introduced mastery nodes, to reflect the idea of granting skill mastery for each skill based on all the skill combinations' knowledge levels. The study reported positive results compared to the vanilla model.

To model interactions between skills, Khajah et al. (2016) proposed an approach to automatic skill discovery that decides on collapsing of different questions to form a single skill. For example, if two skills are highly similar or overlapping, such that learning one predicts learning the other, the research suggested treating this sequence as a single skill and training a single BKT instantiation on both trials. Combined with the forgetting and latent abilities, the proposed model achieved results similar to the neural network-based model.

Wang et al. (2016) investigated the BKT application in the MOOC environment. Based on the chapters and related videos, the proposed model assumed a multi-grained BKT model. When a student masters coarse-grained KC, the probability of mastering one of its fine-grained KCs depends on the probability considering the average correct rate of relevant questions. The study reported improved results over the vanilla model.

Meng et al. (2019) examined the relationship between different knowledge nodes by assuming that the understanding of a certain skill will also deepen, to some extent, the understanding of other skills. Utilizing the parametric matrix of skills initially defined by experts, the gradient descent technique allowed the training of the parameters according to the student's answer data. The proposed model achieved higher prediction accuracy than the vanilla model.

If the vanilla BKT simultaneously traced the cognitive states of students' multiple skills, its time complexity would increase exponentially with the number of skills. Sun et al. (2022) proposed a genetic algorithm and the multi-skills BKT that handles multiple skills simultaneously. Experiments on real datasets showed that the model significantly improved prediction performance over the BKT.

The recent research of Chan et al. (2022) proposed Corrigible Knowledge Tracing (CKT) that assumes students can learn from mistakes in answering multi-skill questions. The experiment results showed that the performance of the proposed model is on par with state-of-the-art approaches.

Hawkins and Heffernan (2014) investigated the relations between questions. Their model investigated how the similarity between the question the student is currently working on and the one the student worked on just before it impacts the performance. The additional node in the network architecture depends on whether the same template generates the previous question as the current one. The results revealed moderate improvement over the vanilla model.

Along with the already mentioned research (Wang et al. 2016), Machardy and Pardos investigated the application of BKT to the MOOC (2015). Exploring the inclusion or omission of video activity, the study investigated the relevance of video lectures to associated assessments. The study demonstrated that video observations can offer information relevant to student behaviour prediction.

6.1.8 Question difficulty (EA09)

Although first investigated in 2010 as part of the scaffolding process by the system (Wang et al. 2010; Wang & Heffernan 2013), equal question difficulty assumption proved to be a significant drawback of the vanilla model. Penalizing the overall score when a student used the hint or multiple attempts to answer the question indirectly led to the question difficulty.

The introduction of the question difficulty into the vanilla model was part of several enhanced models (Baker et al. 2018; David et al. 2016; González-Brenes et al. 2014; Khajah et al., 2014, b; Ostrow et al. 2015; Pardos et al. 2013; Pardos & Heffernan 2011; Zhou et al. 2017).

The Item difficulty effect model (KT-IDEM) (Pardos et al. 2013; Pardos & Heffernan 2011) and the BKT model for periodic learning (Baker et al. 2018) introduced various difficulties by individualizing guess and slip parameters per each question.

The latent factors model introduced by Khajah et al. (2014b) investigated the estimation of BKT parameters per skill, question, and student, while the FAST-BKT model (Gonzalez-Brenes et al. 2014) substituted the logistic regression with coefficients for each skill and each item.

The further research of Khajah and coauthors (2014) and the model proposed by Zhou et al. (2017) incorporated the Item Response Theory (IRT) to deal with question difficulty.

Ostrow et al. (2015) used the per cent correct measure to define the question difficulty, unlike David et al. (2016), who used a weighted score per question and sequenced the content based on the enhanced BKT model.

6.1.9 Multiple attempts (EA10)

While the vanilla model takes into account only the first answer attempt, several publications proposed improvements on this assumption by taking into account student's multiple attempts (Bhatt et al. 2020; Gonzalez-Brenes et al. 2014; Pardos et al. 2013; Yudelson et al. 2008).

Modelling all answer attempts in the ITS environments was proposed by Yudelson et al. (2008) and Gonzales-Brenes et al. (2014).

Pardos et al. (2013) enabled multiple answer attempts in the MOOC platform indicating it is important for the knowledge modelling process. Students were allowed unlimited answer attempts in homework assignments, but only three in midterm and final exam.

Bhatt et al. (2020) examined multiple attempts in the BKT model for the online learning platform in human resources training. Open navigation online learning system allowed learners to choose the next learning activity. The study investigated whether predicting proficiency and communicating it to learners can save time for learners within a course. The proposed model resulted in the more accurate proficiency prediction.

The indirect use of the exact number of attempts in the context of the partial score per question was also part of the BKT research (David et al. 2016; Ostrow et al. 2015; Wang et al. 2010; Wang & Heffernan 2013).

6.1.10 Student characteristics (EA11)

Student characteristics were used in 20 enhanced BKT models, making it the most frequent aspect of BKT enhancements. Most of the research incorporated individual features of various student learning abilities (Agarwal et al. 2018; Baker et al. 2008a, 2010, 2008b; Eagle et al. 2018; Gorgun & Bulut 2022; Khajah et al. 2016; Khajah, Wing, et al., 2014; Nedungadi & Remya 2014; Yudelson 2021; Zhu et al. 2018).

Baker and coauthors (Agarwal et al. 2018; Baker et al. 2008a, 2010, 2008b) proposed already mentioned contextual estimation of guess and slip probabilities based on action–learning speed and history of help-seeking.

To account for individual variation in student ability, the extended BKT models modulated the slip and guess probabilities by a latent ability parameter inferred from the data (Khajah et al. 2016; Khajah et al., 2014a). The models assumed that students with stronger abilities have a lower slip and higher guess probabilities. When the models interacted with new students, they initially used the posterior predictive distribution on abilities. As they observed responses from the new student, uncertainty in the student's ability diminished, yielding better predictions for the student.

The model by Nedungadi and Remya (2014) aimed to improve student performance prediction using an enhanced BKT model with individual priors for each student and skill, and dynamic clustering of students based on changing learning ability.

Eagle et al. (2018) provided individualization by proposing the model fit per student and revealed that we can effectively predict the student parameters for the new lesson by using features derived from prior lessons, and before tutor text-reading data.

Zhu et al. (2018) proposed a model that incorporated the detection of the cognitive inflection points where examinees' cognition levels change a lot, and then integrated that information into a refined BKT model to quantify the knowledge state accurately.

The recent research by Yudelson (2021) focused on the individualization of BKT using a mechanism of an Elo rating schema initially used to rate chess players.

While the focus of the recent research by Gorgun and Bulut (2022) was to analyse the improvement of prediction accuracies of BKT models, the researchers did not include the disengaged student behaviours found in data (i.e. hint abusers and rapid guessers). By dealing with this specific noise in data, the prediction accuracy of BKT models substantially increased.

Other student characteristics also included mental state measures (EEG) (Xu et al. 2014), affective states (Corrigan et al. 2015; Spaulding et al. 2016), eye-tracking data (Rau & Pardos 2016), response time (Lin et al. 2016; Lin & Chi 2016), and brain activity measures (fMRI) (Halpern et al. 2018).

Besides changes in the network architecture, the research on the network parameters allowed the individualization of the BKT model (Khajah et al., 2014a; Lee & Brunskill 2012; Pardos et al. 2012; Wang & Heffernan 2012).

6.1.11 Tutor interventions (EA12)

Tutor interventions as help or hints were also frequently investigated enhancement aspect of the BKT model (Agarwal et al. 2018; Baker et al. 2008a, 2010, 2008b; 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; Wang et al. 2010; Wang & Heffernan 2013; Yudelson et al. 2008).

Yudelson (2008) proposed a model that treats hints differently from mistakes, including the hint request as the third observable performance state.

The Help BKT model (Beck et al. 2008; Chang et al. 2006b) considered two effects of tutor help, including the scaffolding of immediate performance and persistent learning/knowledge.

Contextual guess and slip method used information on help-seeking (Agarwal et al. 2018; Baker et al. 2008a, 2010, 2008b).

The models proposed by Ostrow et al. (2015), and Wang and coauthors (Wang et al. 2010; Wang & Heffernan 2012) incorporated hints.

The intervention BKT model (Lin et al. 2016; Lin & Chi 2016) incorporated instructional interventions by the system and the student response time into the enhanced model.

The model with eye-tracking data (Rau & Pardos 2016) also assumed hints.

Schodde et al. (2017) investigated the impact of tutoring actions on the subsequent observations and skills in the environment of a child-robot language tutor.

6.1.12 Noise in data (EA13)

In addition to the student characteristics and tutor interventions, as enhancements not assumed in the vanilla model, the noise in data was also part of the further research (Beck & Sison 2004; Falakmasir et al. 2015; Gorgun & Bulut 2022).

Beck and Sison (2004) investigated the automated speech recognition domain and proposed the enhanced BKT model to treat noisy data.

Falakmasir et al. (2015) proposed a model that used feature compensation and model compensation paradigms to conceptualize a more flexible and robust BKT model. The proposed model used 3-grams of the binary sequences of student answers to deal with the noise in the observations.

Gorgun and Bulut (2022) investigated the disengaged student behaviour (i.e. hint abusers and rapid guessers) as the noise in data. During the data pre-processing phase, the researchers removed these behaviours. The results showed an increase in the prediction accuracy of the examined models.

6.1.13 Computational methods (RQ1)

The increased interest in the research of the expectation–maximization method resulted in the detection of difficulties.

Although firstly reported as the model identifiability problem (Beck & Chang 2007), Doroudi and Brunskill (2017) revealed that under mild conditions on the parameters, the BKT model is identifiable and it 'only' suffers from the local optima problem. The recent open-source accessible and computationally efficient Python library of BKT models—pyBKT (Badrinath et al. 2021), also includes the expectation–maximization method to fit the BKT parameters. Since the expectation–maximization algorithm is susceptible to converging to the 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.

As reported, the expectation–maximization method is prone to semantic model degeneracy and can be inconsistent with the conceptual assumptions underlying the BKT model (Doroudi & Brunskill, 2017).

6.2 Evaluation approaches (RQ2)

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

Regarding the increased use of the Cognitive Tutor and ASSISTments-based datasets, it is important to note these were the only publicly available datasets identified in the review. They were part of the Educational Data Mining KDD Cup Challenge, hosted by PSLC DataShop (2010), which highly contributed to the awareness of the importance of replicability and comparison of the proposed models.

By using the typical metrics for classification tasks in the machine learning field, 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 literature frequently reported on the first type of evaluation and performance measures as RMSE, AUC–ROC and accuracy. In terms of 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 (Baker et al. 2010; Beck & Sison 2004; Lin et al. 2016; Lin & Chi 2016; Schodde et al. 2017; Yudelson et al. 2008). In the latter case, the correlation was a frequent model performance measure. The researchers typically compared the proposed models to the vanilla BKT model. While most of the enhanced BKT models reported better results than the vanilla model, there is research that reported unexpected and mixed results for both types of evaluation approaches. The discussion of the previous results in the following paragraphs also represents the methodological note to the further research of BKT enhancements.

The unexpected performance results as the model drawbacks were reported in the case of the Help BKT model (Beck et al. 2008; Chang et al. 2006a), the Affective BKT model (Corrigan et al. 2015), the BKT with the eye-tracking model (Rau & Pardos 2016) and the model by Adjei et al. (2013). 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 the 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) (Baker et al. 2008a, 2008b), the Item Difficulty Effect Model (KT-IDEM) (Pardos & Heffernan 2011), the Student Skill BKT model (SS-BKT) (Wang & Heffernan 2012) and the BKT with tutoring actions (Schodde et al. 2017). Besides predicting in-tutor performance, the CGS-BKT investigated the student performance on the post-test outside the system. 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, 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 significantly differ from the vanilla model in the post-test results.

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 (Chang et al. 2006a; Murphy 2001), the hmm-scalable implementation (Yudelson et al. 2013) and the approach proposed for MOOC resources by Xu et al. (2015). The recent positive example is the pyBKT library (Badrinath et al. 2021), an accessible and computationally efficient BKT implementation framework.

7 Conclusion

It has been Twenty-five Years of BKT research, and the vanilla model is still a representative probabilistic 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 only 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 improved models. The most extensive student modelling review focused on ML approaches and encompassed 18 student models based on Bayesian networks (Liu et al. 2021b). The other review study elaborated on the 8 BKT models from 2010 to 2015 (Sani et al. 2016). The systematic review from 2013 discussed the 13 ML student models and 18 Bayesian network-based models (Chrysafiadi & Virvou 2013). In addition, each publication that proposes an enhanced BKT model also provides only a subset of the background literature because of the limited format and different purposes. 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 a unique set of criteria, including 10 enhancement aspects based on the vanilla model assumptions and 3 ones 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 commonly investigated enhancement aspects were the domain knowledge properties (assumed as independent knowledge components in the vanilla model) and the question difficulty (taken as constant in the vanilla model), representing the educational context-based enhancements. Although less investigated in the literature, the obvious drawbacks of the vanilla model referred to the binary states and exclusion. As suggested in the reference literature, the uncertainty and fuzziness in this context would be more appropriate. Besides the previous, 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 enhancement aspects can be differed as generally applicable and dependent on the capabilities of educational settings. While most enhancement aspects already defined by the vanilla model represent generally applied enhancements, if the educational environment enables specific features (e.g. student characteristics and tutor interventions), the research suggests checking their contribution to the knowledge inference process.

Besides the previous aspects, the BKT models were reviewed regarding the incorporated computational methods. The enhanced models generally improved the expert-based estimations of BKT parameters per skill 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.

Various ITSs and MOOCs represent a broad testing ground for future research, and standard performance measures ensure future comparisons of the results. While the enhanced models generally included the prediction of in-tutor performance as the evaluation approach, future research should focus on knowledge mastery because the duration of the teaching and learning process depends on it. The literature suggests using additional post-tests to examine the same domain knowledge using questions similar to those in the examined educational platform.

The major limitation in the BKT research is model availability. A positive example in this context is the pyBKT library (Badrinath et al. 2021), 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.

This systematic review of the BKT enhancements is unbiased, exhaustive, and inspiring for future research.

Author contributions ISG had the idea for the article, performed the literature search and data analysis, and drafted the first version of the manuscript. All authors critically revised the work and approved its final version.

Funding This work has been supported by the Office of Naval Research under the grant N00014-20-1-2066 Enhancing Adaptive Courseware based on Natural Language Processing.

Declarations

Competing interests The authors declare that they have no competing interests (appropriate disclosures in the further text).

Appendix A

See Table 12

Table 12 Enhanced BKT models

#	Publications	Year	Model name	Short description	Enhancements (E01-E13)
1	Beck and Sisson (2004)	2004	BKT for noisy data	The model enhanced to treat noisy data in the automated speech recognition domain	E13
2	Chang et al. (2006b); Beck et al. (2008)	2006	Help BKT	The model considers two effects of the tutor's help: scaffolding immediate performance and persistent learning/knowledge	E07, E12
3	Beck (2007); Beck and Chang (2007)	2007	BKT with the Dirichlet Priors method	The Dirichlet Priors method chooses the most similar parameter estimates among different skills	-
4	Yudelson et al. (2008)	2008	17 BKT models	17 BKT models treat hints differently from mistakes, consider partially mastered skills, forgetting, and all attempts made by a student	E01, E02, E07, E10, E12
5	Baker et al. (2008b); Baker et al. (2008a)	2008	BKT with Contextual Guess and Slip method (CGS-BKT)	The Contextual Guess and Slip method (CGS-BKT) is based on the dynamic estimation of parameters using the contextual information of the action–learning speed and history of help-seeking	E04, E05, E11, E12
6	Ritter et al. (2009)	2009	BKT with clustered parameters	The method with clustered parameters investigates the reduction of the parameter space	-

Table 12 (continued)

#	Publications	Year	Model name	Short description	Enhancements (E01-E13)
7	Baker et al. (2010)	2010	BKT with Contextual Guess and Slip and Brute force methods (CGS-BKT-BF)	The Contextual Guess and Slip and Brute force methods (CGS-BKT-BF) investigate the model performance on the post-test outside of the tutor	E04, E05, E11, E12
8	Pardos and Heffernan (2010a)	2010	Prior Per Student BKT (PPS BKT)	The model focuses on individualizing the prior knowledge parameter by comparing the random values, cold start heuristic, and per cent correct heuristic approaches	E03
9	Pardos and Heffernan (2010b)	2010	BKT with the bounded sum of guess and slip	The model that investigates the Expectation–Maximization parameter space and the initial values of guess and slip parameters	-
10	Wang et al. (2010); Wang and Heffernan (2013)	2010	BKT with Partial Credit	The model incorporates partial credit scores using the hints and attempts penalties	E02, E10, E12
11	Qiu et al. (2011)	2011	KT-Forget and KT-Slip	The models focus on the poor student performance on the next day of learning, assuming it as forgetting or slipping	E05, E07
12	Pardos and Heffernan (2011)	2011	Item Difficulty Effect Model (KT-IDEM)	The model focuses on the item difficulty by individualizing guess and slip parameters per each question	E04, E05, E09
13	Wang and Heffernan (2012)	2012	Student Skill BKT (SS-BKT)	The model focuses on the simultaneous parameter estimation per skill and per student	E11

Table 12 (continued)

#	Publications	Year	Model name	Short description	Enhancements (E01-E13)
14	Lee and Brunskill (2012)	2012	BKT with a 4-parameter fit per student	The model investigates the individualization per student using all 4 parameters and the expected number of practice opportunities for a student to reach mastery	E11
15	Pardos et al. (2012)	2012	Clustered BKT	The model investigates the individualization per student group by clustering students based on their performance and using different parameters for students in different clusters	E11
16	Falakmasir et al. (2013)	2013	BKT with spectral parameter fitting	The model investigates a spectral method for learning the parameters of BKT directly from students' sequences of correct/incorrect responses	-
17	Pardos et al. (2013)	2013	Count model, Item Difficulty Model (IDEM), and IDEM Count	The models adapt to the MOOC environment and its various resources; the Count model allows multiple attempts, the IDEM model investigates the variation between questions within a knowledge component, and IDEM Count investigates the combination of both	E09, E10
18	Sao Pedro et al. (2013)	2013	Scaffolding and Topic BKT	The model focuses on the system scaffolding and changing science topics by adding them as observable elements of the HMM	E08, E06

Table 12 (continued)

#	Publications	Year	Model name	Short description	Enhancements (E01-E13)
19	Yudelson et al. (2013)	2013	Individualized BKT (iBKT)	The model investigates the introduction of individualized parameters of initial mastery probabilities and skill learning probabilities in an incremental manner	E03, E06
20	Adjei et al. (2013)	2013	BKT with two learning rates	The model with two learning rates	E06
21	Xu and Mostow (2013)	2013	BKT with Item Response Theory IRT—2013	The model that incorporates Item Response Theory	E03
22	Sao Pedro et al. (2014)	2014	Partial Skill Transfer BKT (PST-BKT)	The model aims to capture the possibility of a partial transfer of a skill	E08
23	Hawkins et al. (2014)	2014	BKT with Empirical Probabilities method (BKT-EP)	The model investigates the method that sacrifices the precision of optimization techniques for the efficiency and interpretability of empirical estimation	-
24	Hawkins and Heffernan (2014)	2014	BKT Same Template (BKT-ST)	The model considers the similarity between the problem the student is currently working on and the one the student worked on just before it	E08
25	Xu et al. (2014)	2014	Electroencephalography BKT (Binary EEG-BKT and Multi-dimensional continuous EEG-LRKT) models	The models incorporate the mental state measures using EEG	E11
26	Nedungadi et al. (2014)	2014	Personalized Clustered BKT (PC-BKT)	The model focuses on incorporating the individual priors for each student and skill, and dynamic clustering of students based on changing learning ability	E03, E11

Table 12 (continued)

#	Publications	Year	Model name	Short description	Enhancements (E01-E13)
27	Gonzales-Brenes et al. (2014)	2014	Feature Aware Student Knowledge Tracing BKT (FAST-BKT)	The model proposes additional features for BKT and deals with Multiple Subskills, Temporal Item Response Theory (IRT), and Expert Knowledge	E08, E09, E10
28	Khajah et al. (2014b)	2014	Latent-factor Knowledge Tracing (LF-KT)	The model incorporates the temporal sequence of experience (BKT) and individual differences in student ability and problem difficulty (LFM)	E09, E11
29	Khajah et al. (2014a)	2014	BKT with Item Response Theory IRT—2014	The model combines BKT and Item Response Theory, which infers individual differences amongst students and questions	E09, E11
30	Machardy and Pardos (2015); Machardy (2015)	2015	Template Videos BKT and Template-1-Video BKT models	The models are adapted to the MOOC environment and focus on incorporating the usage of video content in addition to assessment activity	E08
31	Corrigan et al. (2015)	2015	Affective BKT	The model considers the affective states of a student (confusion, frustration, boredom, concentration)	E11
32	Lee et al. (2015); Gweon et al. (2015)	2015	BKT with Monte Carlo method	The model investigates the Monte Carlo method in parameter fitting procedure and clusters knowledge emergence space defined by guess, slip, and transition parameters	-

Table 12 (continued)

#	Publications	Year	Model name	Short description	Enhancements (E01-E13)
33	Nedungadi and Remya (2015)	2015	Personalized Clustered BKT with Forgetting (PC-BKT-F) and Decay (PC-BKT-D)	The models incorporate forgetting into the PC-BKT model	E03, E07
34	Song et al. (2015)	2015	BKT with Prediction of First-encounter Knowledge (PSFK)	The model focuses on the first-encounter knowledge prediction using user-based collaborative filtering	E03
35	Falakmasir et al. (2015)	2015	Spectral BKT	The model uses spectral observations – 3-g of the binary sequences of student answers to deal with the noise in the observations	E13
36	Ostrow et al. (2015)	2015	BKT with partial credit score and problem difficulty	The model addresses the partial credit score based on the first response, attempt count, and hint usage and problem difficulty issue based on per cent correct	E02, E09, E10, E12
37	Rau and Pardos (2016)	2016	BKT with Eye-tracking data	The model incorporates the visual attention features of areas of interest (problem-solving, hints, and periodic table)	E11, E12
38	Spaulding et al. (2016)	2016	Affect-Aware BKT	The model incorporates automatic detection of visual expressions (smiling, engaging)	E11

Table 12 (continued)

#	Publications	Year	Model name	Short description	Enhancements (E01-E13)
39	Eagle et al. (2016a,); Eagle et al. (2016b); Eagle et al. (2017)	2016	BKT fit per student pre-activities	The model is based on performance in activities that naturally precede tutor modules: reading online instructional text and taking a conceptual knowledge pre-test to predict performance in the first two practice opportunities	E03
40	Khajah et al. (2016)	2016	BKT with forgetting, skill discovery, and latent student-abilities (BKT + FSA)	The model incorporates dealing with forgetting, automatic skill discovery, and latent student-abilities	E07, E08, E11
41	Lin and Chi (2016); Lin et al. (2016)	2016	Intervention BKT	The model incorporates the student response time and instructional interventions by the system	E11, E12
42	Yudelson (2016)	2016	Individualized BKT (iBKT) – Hierarchical Bayesian Models (HBM)	The model investigates individualization using HBM and in addition to capturing student-level variability in the data weighs the contribution of per-student and per-skill effects to the overall variance in the data	-
43	David et al. (2016)	2016	BKT with partial credit scores, item difficulty, and multiple attempts	The model investigates the partial credit scores implemented to previously proposed enhanced models (KT-IDEM and Count IDEM)	E02, E09, E10
44	Wang et al. (2016)	2016	Multi-grained BKT and Historical BKT	The models investigate the relations between subskills in terms of the hierarchical properties (the structure) and temporal properties in the MOOC environment	E02, E08

Table 12 (continued)

#	Publications	Year	Model name	Short description	Enhancements (E01-E13)
45	Huang and Brusilovky (2016); Huang et al. (2016)	2016	BKT for skill combinations	The models that investigate skill combinations	E08
46	Schodde et al. (2017)	2017	BKT with tutoring actions	The model investigates the impact of tutoring actions on the subsequent observations and skills in the environment of a child-robot language tutor	E01, E12
47	Wang et al. (2017)	2017	BKT with Item Response Theory-2PL (BKT-IRT-2PL)	The model incorporates Item Response Theory (2PL) model that estimates students' initial knowledge status and joins it with the discrimination and difficulty of each skill estimated to evaluate the probability of knowing a skill before training it	E03
48	Zhou et al. (2017)	2017	BKT for multiple subskills	The model that includes the guessing and slipping probabilities based on the student ability of each subskill	E04, E05, E08, E09
49	Zhang and Yao (2018)	2018	Three learning states BKT (TLS-BKT)	The model proposes the unlearned state, the learning state, and the learned state and includes the trisection of the learning process by using three-way decision theory	E01
50	Agarwal et al. (2018)	2018	BKT with stable parameters	The model proposes a new iterative method based on contextual guess and slip estimation that converges to stable estimates for skill-level guess and slip parameters	E04, E05, E11, E12

Table 12 (continued)

#	Publications	Year	Model name	Short description	Enhancements (E01-E13)
51	Zhu et al. (2018)	2018	BKT with Temporal Difference (TD-BKT)	The model incorporates temporal difference information by detecting cognitive inflection points where the examinee's cognitive level changes a lot in the MOOC environment	E11
52	Halpern et al. (2018)	2018	BKT with continuous functional magnetic resonance imaging signals (fMRI-BKT)	The model focuses on integrating neuroimaging recordings (fMRI) into BKT and proposes Study and Decay probabilities as groups of transitions	E01, E07, E11
53	Baker et al. (2018)	2018	BKT with Periodic Learning (BKT-PL)	The model investigates the assumption that most learning occurs during the use of declarative content rather than between exercises and updates knowledge probability after the set of problems instead of each problem	E06, E09
54	Eagle et al. (2018)	2018	BKT fit per student pre-activities (no pre-test measure)	The model uses data collected from student reading and the performance on the first two tutor lessons to predict individualized difference weights of BKT parameters for the third lesson	E11
55	Pu et al. (2019)	2019	BKT with parallel Expectation–Maximization method	The model proposes a new parallel BKT open-source tool based on the Spark computational framework with the method of automatic tuning of initial parameters	-
56	Meng et al. (2019)	2019	Cross-Skill BKT (CS-BKT)	The model considers the relationship between skills (HMM knowledge nodes)	E08

Table 12 (continued)

#	Publications	Year	Model name	Short description	Enhancements (E01-E13)
57	Bhatt et al. (2020)	2020	Attempt enhanced BKT	The model incorporates multiple attempts and proposes the fitting of parameters per attempt (up to four attempts) into an open-navigation system for adult workforce skill development	E10
58	Liu et al. (2021a,)	2021	Fuzzy BKT (FBKT) and Type-2 Fuzzy theory BKT (T2FBKT)	The models are based on type-1 and type-2 fuzzy theory to deal with vanilla BKT assumptions of binary states of knowledge and performance nodes	E01, E02
59	Yudelson (2021)	2021	Elo based iBKT	The model focuses on the individualization of BKT using a mechanism of an Elo rating schema	E11
60	Gorgun and Bulut (2022)	2022	BKT without Disengaged behaviour	The model is based on the pre-processed dataset that does not include disengaged student behaviour	E11, E13
61	Sun et al. (2022)	2022	Multi-skills BKT	The model uses a genetic algorithm to propose the multi-skills BKT	E08
62	Chan et al. (2022)	2022	Corrigible Knowledge Tracing (CKT)	The model assumes students can learn from mistakes in answering multi-skill questions	E08

References

Adjei, S., Salehizadeh, S., Wang, Y., & Heffernan, N. T.: Do students really learn an equal amount independent of whether they get an item correct or wrong?. In: D'Mello, S. K., Calvo, R. A., & Olney A. (eds.) Proceedings of the 6th International Conference on Educational Data Mining, Memphis, Tennessee, USA, pp. 304–305. International educational data mining society (2013). https://dblp.org/rec/conf/edm/AdjeiSWH13.html

Agarwal, D., Babel, N., & Baker, R. S.: Contextual derivation of stable BKT parameters for analysing content efficacy. In: Boyer K.E., & Yudelson M. (eds.) Proceedings of the 11th International Conference on Educational Data Mining, EDM 2018, Buffalo, NY, USA. International Educational Data Mining Society (IEDMS) (2018). https://dblp.org/rec/conf/edm/AgarwalBB18.html

Anderson, J. R.: Rules of the mind (pp. ix, 320). Lawrence Erlbaum Associates (1993)

Anouar Tadlaoui, M., Souhaib, A., Khaldi, M., Carvalho, R.: Learner modeling in adaptive educational systems: a comparative study. Int. J. Modern Educ. Comput. Sci. **8**, 1–10 (2016). https://doi.org/10.5815/ijmecs.2016.03.01

Atkinson, R.C.: Optimizing the learning of a second-language vocabulary. J. Exp. Psychol. **96**(1), 124–129 (1972). https://doi.org/10.1037/h0033475

Badrinath, A., Wang, F., & Pardos, Z. A.: pyBKT: an accessible python library of Bayesian knowledge tracing models. CoRR (2021) https://arxiv.org/abs/2105.00385

Baker, R.S., Inventado, P.S.: Educational data mining and learning analytics. In: Larusson, J.A., White, B. (eds.) Learning analytics: from research to practice, pp. 61–75. Springer, Berlin (2014). https://doi.org/10.1007/978-1-4614-3305-7_4

Baker, R. S., Gowda, S. M., & Salamin, E.: 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 Vol 8 (2018)

Baker, R. S., Corbett, A. T., & Aleven, V. (2008a). More accurate student modeling through contextual estimation of slip and guess probabilities in bayesian knowledge tracing. In: Woolf, B. P., Aïmeur, E., Nkambou, R., Lajoie S. P. (eds.) Proceedings of the 9th Inernational Conference on Intelligent Tutoring Systems, ITS 2008, Montreal, Canada Vol. 5091, pp. 406–415. Springer (2008). https://doi.org/10.1007/978-3-540-69132-7_44

Baker, R. S., Corbett, A. T., & Aleven, V.: Improving contextual models of guessing and slipping with a truncated training set. In: International Conference on Educational Data Mining, EDM (2008b). https://doi.org/10.1184/R1/6470135.v1

Baker, R. S., Corbett, A. T., Gowda, S. M., Wagner, A. Z., MacLaren, B. A., Kauffman, L. R., Mitchell, A. P., & Giguere, S.: Contextual slip and prediction of student performance after use of an intelligent tutor. In: P. Bra, A. Kobsa, & D. N. Chin (eds.) User Modeling, Adaptation, and Personalization, 18th International Conference, UMAP 2010, Big Island, HI, USA, June 20–24, 2010. Proceedings Vol. 6075, pp. 52–63. Springer (2010). https://doi.org/10.1007/978-3-642-13470-8_7

Beck, J. E., & Chang, K.: Identifiability: a fundamental problem of student modeling. In: Conati, C., McCoy, K. F., & Paliouras G., (eds.) User Modeling 2007, 11th International Conference, UM 2007, Corfu, Greece, June 25–29, 2007, Proceedings Vol. 4511, pp. 137–146. Springer (2007). https://doi.org/10.1007/978-3-540-73078-1_17

Beck, J. E., & Sison, J.: Using knowledge tracing to measure student reading proficiencies. In: Lester, J. C., Vicari, R. M., Paraguaçu F. (eds.) Intelligent Tutoring Systems, 7th International Conference, ITS 2004, Maceiò, Alagoas, Brazil, August 30—September 3, 2004, Proceedings Vol. 3220, pp. 624–634. Springer (2004). https://doi.org/10.1007/978-3-540-30139-4_59

Beck, J. E., Chang, K., Mostow, J., & Corbett, A. T.: Does help help? introducing the bayesian evaluation and assessment methodology. In: Woolf, B. P., Aïmeur, E., Nkambou, R. & Lajoie S. P. (eds.) Intelligent Tutoring Systems, 9th International Conference, ITS 2008, Montreal, Canada, June 23–27, 2008, Proceedings Vol. 5091, pp. 383–394. Springer (2008). https://doi.org/10.1007/978-3-540-69132-7_42

Beck, J. E.: Difficulties in inferring student knowledge from observations (and why you should care) (2007)

Bhatt, S. P., Zhao, J., Thille, C., Zimmaro, D., & Gattani, N.: Evaluating Bayesian knowledge tracing for estimating learner proficiency and guiding learner behavior. In: Joyner, D., Kizilcec, R. F., & Singer S. (eds.) L@S'20: Seventh ACM Conference on Learning @ Scale, Virtual Event, USA, August 12–14, 2020 (pp. 357–360). ACM (2020). https://doi.org/10.1145/3386527.3406746

Brusilovsky, P., Millán, E.: User models for adaptive hypermedia and adaptive educational systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The adaptive web: methods and strategies of web personalization, pp. 3–53. Springer, Berlin (2007)
Cen, H., Koedinger, K., Junker, B.: Learning factors analysis–A general method for cognitive model evaluation and improvement. In: Ikeda, M., Ashley, K.D., Chan, T.-W. (eds.) Intelligent Tutoring Systems, pp. 164–175. Springer, Berlin (2006). https://doi.org/10.1007/11774303_17
Chan, K. I., Tse, R., & Lei, P. I. S.: Tracing students' learning performance on multiple skills using Bayesian methods. In: Proceedings of the 6th International Conference on Education and Multimedia Technology, pp 84–89. (2022) https://doi.org/10.1145/3551708.3556202
Chang, K., Beck, J. E., Mostow, J., & Corbett, A. T.: (2006a). 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, pp 104–113.
Chang, K., Beck, J. E., Mostow, J., & Corbett, A. T. (2006b). Does help help? a bayes net approach to modeling tutor interventions. In: AAAI2006 Workshop on Educational Data Mining, Boston, Massachusetts.
Chrysafiadi, K., Virvou, M.: Student modeling approaches: a literature review for the last decade. Expert Syst. Appl. **40**(11), 4715–4729 (2013). https://doi.org/10.1016/j.eswa.2013.02.007
Corbett, A.T., Anderson, J.R.: Knowledge tracing: modeling the acquisition of procedural knowledge. User Model. User-Adap. Inter. **4**(4), 253–278 (1995)
Corrigan, S., Barkley, T., & Pardos, Z. A.: Dynamic approaches to modeling student affect and its changing role in learning and performance. In: Ricci, F., Bontcheva, K., Conlan, O., Lawless, S. (eds.) User modeling, adaptation and personalization—23rd international conference, UMAP 2015, Dublin, Ireland. Proceedings Vol. 9146, pp. 92–103. Springer (2015). https://doi.org/10.1007/978-3-319-20267-9_8
David, Y. B., Segal, A., & Gal, Y.: Sequencing educational content in classrooms using Bayesian knowledge tracing. In: Gasevic, D., Lynch, G., Dawson, S., Drachsler, H., & Rosé C. P. (eds.) Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, LAK 2016, Edinburgh, United Kingdom pp. 354–363. ACM (2016). https://doi.org/10.1145/2883851.2883885
DeFalco, J.A., Sinatra, A.M.: Adaptive instructional systems: the evolution of hybrid cognitive tools and tutoring systems. In: Sottilare, R.A., Schwarz, J. (eds.) Adaptive Instructional Systems, pp. 52–61. Springer, Berlin (2019). https://doi.org/10.1007/978-3-030-22341-0_5
Desmarais, M.C., Baker, R.S.: A review of recent advances in learner and skill modeling in intelligent learning environments. User Model. User-Adap. Inter. **22**(1), 9–38 (2012). https://doi.org/10.1007/s11257-011-9106-8
Doroudi, B.: (2017). https://dblp.org/rec/conf/edm/DoroudiB17.html
Eagle, M., Corbett, A. T., Stamper, J. C., McLaren, B. M., Baker, R. S., Wagner, A. Z., MacLaren, B. A., & Mitchell, A. P.: Predicting individual differences for learner modeling in intelligent tutors from previous learner activities. In: Vassileva, J., Blustein, J., Aroyo, L., & D'Mello S. K. (eds.) Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization, UMAP 2016, Halifax, NS, Canada, pp. 55–63. ACM (2016). https://doi.org/10.1145/2930238.2930255
Eagle, M., Corbett, A. T., Stamper, J. C., McLaren, B. M., Wagner, A. Z., MacLaren, B. A., & Mitchell, A. P.: Estimating individual differences for student modeling in intelligent tutors from reading and pretest data. In: Micarelli, A., Stamper, J. C., & Panourgia K. (eds.) Intelligent tutoring systems—13th International Conference, ITS 2016, Zagreb, Croatia, June 7–10, 2016. Proceedings Vol. 9684, pp. 133–143. Springer (2016). https://doi.org/10.1007/978-3-319-39583-8_13
Eagle, M., Corbett, A. T., Stamper, J. C., McLaren, B. M., Baker, R. S., Wagner, A. Z., MacLaren, B. A., & Mitchell, A. P.: Exploring learner model differences between students. In: André, E., Baker, R. S., Hu, X., Rodrigo, M. M. T., & Du Boulay B. (eds.) Artificial Intelligence in Education—18th International Conference, AIED 2017, Wuhan, China, June 28—July 1, 2017, Proceedings (Vol. 10331, pp. 494–497). Springer (2017). https://doi.org/10.1007/978-3-319-61425-0_48
Eagle, M., Corbett, A. T., Stamper, J. C., & McLaren, B. M.: Predicting Individualized Learner Models Across Tutor Lessons. In: Boyer, K.E., Yudelson M. (eds.) Proceedings of the 11th International Conference on Educational Data Mining, EDM 2018, Buffalo, NY, USA. International Educational Data Mining Society (IEDMS) (2018). https://dblp.org/rec/conf/edm/EagleCSM18.html
Falakmasir et al. (2013) https://dblp.org/rec/conf/edm/FalakmasirPGB13.html
Falakmasir, M. H., Yudelson, M., Ritter, S., & Koedinger, K. R.: Spectral Bayesian knowledge tracing. In: Santos, O. C., Boticario, J., Romero, C., Pechenizkiy, M., Merceron, A., Mitros, P., Luna, J. M.,

Mihaescu, M. C., Moreno, P., Hershkovitz, A., Ventura, S., Desmarais M. C. (eds.) Proceedings of the 8th International Conference on Educational Data Mining, EDM 2015, Madrid, Spain, June 26–29, 2015 (pp. 360–363). International Educational Data Mining Society (IEDMS) (2015). https://dblp.org/rec/conf/edm/FalakmasirYRK15.html

Freedman, R., Ali, S.S., McRoy, S.: Links: what is an intelligent tutoring system? Intelligence **11**(3), 15–16 (2000). https://doi.org/10.1145/350752.350756

Gonzalez-Brenes, J., Huang, Y., & Brusilovsky, P.: General features in knowledge tracing to model multiple subskills, temporal item response theory, and expert knowledge. In: The 7th International Conference on Educational Data Mining, pp. 84–91 (2014). https://educationaldatamining.org/EDM2014/uploads/procs2014/long%20papers/84/EDM-2014-Full.pdf

González-Brenes, J. P., Huang, Y., & Brusilovsky, P.: General features in knowledge tracing to model multiple subskills, temporal item response theory, and expert knowledge. In: Stamper, J. C., Pardos, Z. A., Mavrikis, M. & McLaren B. M. (eds.) Proceedings of the 7th International Conference on Educational Data Mining, EDM 2014, London, UK pp. 84–91. International Educational Data Mining Society (IEDMS) (2014). https://dblp.org/rec/conf/edm/HuangGB14.html

Gorgun, G., Bulut, O.: In: Rodrigo, M.M., Matsuda, N., Cristea, A.I., Dimitrova, V. (eds.) Artificial intelligence in education. Posters and late breaking results, workshops and tutorials, industry and innovation tracks, practitioners' and doctoral consortium, pp. 591–594. Springer, Berlin (2022)

Gweon et al. (2015) https://dblp.org/rec/conf/lak/GweonLDTFD15.html

Halpern, D., Tubridy, S., Wang, H. Y., Gasser, C., Popp, P. O., Davachi, L., & Gureckis, T. M.: Knowledge tracing using the brain. In: Boyer K. E., & Yudelson M. (eds.) Proceedings of the 11th International Conference on Educational Data Mining, EDM 2018, Buffalo, NY, USA. International Educational Data Mining Society (IEDMS) (2018). https://dblp.org/rec/conf/edm/HalpernTWGPDG18.html

Harrison, B., & Roberts, D.: A review of student modeling techniques in intelligent tutoring systems. In: Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 8, Article 5 (2012). https://ojs.aaai.org/index.php/AIIDE/article/view/12574

Hawkins, W. J., & Heffernan, N. T.: Using similarity to the previous problem to improve Bayesian knowledge tracing. In: Santos S. G., & Santos O. C., (eds.) 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 Vol. 1183. CEUR-WS.org (2014). https://dblp.org/rec/conf/edm/HawkinsH14.html

Hawkins et al. (2014) https://dblp.org/rec/conf/its/HawkinsHB14.html

Huang, Y., & Brusilovsky, P.: Towards modeling chunks in a knowledge tracing framework for students' deep learning. In: Barnes, T., Chi, M., Feng M. (eds.) Proceedings of the 9th International Conference on Educational Data Mining, EDM 2016, Raleigh, North Carolina, USA, pp. 666–668. International Educational Data Mining Society (IEDMS) (2016).

Huang, Y., Guerra, J., & Brusilovsky, P.: Modeling skill combination patterns for deeper knowledge tracing. In: The 6th International Workshop on Personalization Approaches in Learning Environments (PALE 2016) in the 24th Conf. on User Modeling, Adaptation and Personalization (UMAP 2016) (2016). https://dblp.org/rec/conf/um/HuangGB16.html

Khajah, M., Huang, Y., González-Brenes, J. P., Mozer, M. C., & Brusilovsky, P.: Integrating knowledge tracing and item response theory: a tale of two frameworks. In: CEUR Workshop Proceedings, 1181, 7–15. (2014) https://d-scholarship.pitt.edu/26044/

Khajah, M., Wing, R., Lindsey, R., & Mozer, M.: Incorporating latent factors into knowledge tracing to predict individual differences in learning. In: International Conference on Educational Data Mining, EDM (2014). https://www.semanticscholar.org/paper/Incorporating-Latent-Factors-Into-Knowledge-Tracing-Khajah-Wing/9af7938fd695fd2c90e566923828c6336d3c2292

Khajah, M., Lindsey, R. V., & Mozer, M. (2016). How Deep is Knowledge Tracing? In T. Barnes, M. Chi, & M. Feng (Eds.), *Proceedings of the 9th International Conference on Educational Data Mining, EDM 2016, Raleigh, North Carolina, USA, June 29—July 2, 2016*. International Educational Data Mining Society (IEDMS). https://dblp.org/rec/conf/edm/KhajahLM16.html

Kurup, L. D., Joshi, A., & Shekhokar, N.: A review on student modeling approaches in ITS. In: 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 2513–2517 (2016)

Lee, J. I., & Brunskill, E.: The impact on individualizing student models on necessary practice opportunities. In: Yacef, K., Zaïane, O. R., Hershkovitz, A., Yudelson, M., & Stamper, J. C., (eds.) Proceedings of

the 5th International Conference on Educational Data Mining, Chania, Greece, pp. 118–125. (2012). https://dblp.org/rec/conf/edm/LeeB12.html
Lee et al. (2015) https://dblp.org/rec/conf/lak/LeeGDTFDKPL15.html
Lin, C., Chi, M.: Intervention-BKT: incorporating instructional interventions into bayesian knowledge tracing. In: Micarelli, A., Stamper, J., Panourgia, K. (eds.) Intelligent Tutoring Systems, pp. 208–218. Springer, Berlin (2016)
Lin, C., Shen, S., & Chi, M.: Incorporating student response time and tutor instructional interventions into student modeling. In: Vassileva, J., Blustein, J., Aroyo, L., D'Mello S.K., (eds.) Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization, UMAP 2016, Halifax, NS, Canada pp. 157–161. ACM. (2016) https://doi.org/10.1145/2930238.2930291
Liu, F., Hu, X., Bu, C., Yu, K.: Fuzzy Bayesian knowledge tracing. IEEE Trans. Fuzzy Syst. (2021a). https://doi.org/10.1109/TFUZZ.2021.3083177
Liu, Q., Shen, S., Huang, Z., Chen, E., & Zheng, Y.: A survey of knowledge tracing. https://arxiv.org/abs/2105.15106*[Cs]*. (2021b)
MacHardy, Z., & Pardos, Z. A.: Toward the evaluation of educational videos using bayesian knowledge tracing and big data. In: Kiczales, G., Russell, D. M., Woolf, B.P., (eds) Proceedings of the Second ACM Conference on Learning @ Scale, L@S 2015, Vancouver, BC, Canada, pp. 347–350. ACM (2015) https://doi.org/10.1145/2724660.2728690
MacHardy, Z. (2015). Applications of bayesian knowledge tracing to the curation of educational videos (Technical Report No. UCB/EECS-2015–98, Electrical Engineering and Computer Sciences, University of California at Berkeley). University of California at Berkeley.
Meng, L., Zhang, M., Zhang, W., Chu, Y.: CS-BKT: introducing item relationship to the Bayesian knowledge tracing model. Interact. Learn. Environ. (2019). https://doi.org/10.1080/10494820.2019.1629600
Mitchell, T.M.: Machine learning. McGraw-Hill Education, New York (1997)
Moher et al. (2009) https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000097
Montero, S., Arora, A., Kelly, S., Milne, B., & Mozer, M.: Does deep knowledge tracing model interactions among skills?. In: The 11th international conference on educational data mining, EDM (2018). http://www.scopus.com/inward/record.url?scp=85084011386&partnerID=8YFLogxK
Murphy, K.P.: The bayes net toolbox for MATLAB. Comput. Sci. Stat. **33**, 2001 (2001)
Nedungadi, P., & Remya, M. S.: Predicting students' performance on intelligent tutoring system—Personalized clustered BKT (PC-BKT) model. In: IEEE frontiers in education conference, FIE 2014, Proceedings, Madrid, Spain, pp. 1–6 (2014). https://doi.org/10.1109/FIE.2014.7044200
Nedungadi, P., & Remya, M. S.: Incorporating forgetting in the personalized, clustered, Bayesian knowledge tracing (PC-BKT) model. In: 2015 International Conference on Cognitive Computing and Information Processing (CCIP),pp. 1–5 (2015) https://doi.org/10.1109/CCIP.2015.7100688
Nkambou, R., Bourdeau, J., & Mizoguchi, R. (Eds.) (2010) Advances in Intelligent Tutoring Systems (Vol. 308). Springer, Berlin
Nwana, H.S.: Intelligent tutoring systems: an overview. Artif. Intell. Rev. **4**(4), 251–277 (1990). https://doi.org/10.1007/BF00168958
Ostrow, K., Donnelly, C., Adjei, S., & Heffernan, N. T.: Improving student modeling through partial credit and problem difficulty. In: G. Kiczales, G., Russell, & D.M., Woolf P. (eds.) Proceedings of the Second ACM Conference on Learning @ Scale, L@S 2015, Vancouver, BC, Canada pp. 11–20. ACM (2015). https://doi.org/10.1145/2724660.2724667
PSLC DataShop. (2010). Educational data mining challenge—KDD Cup 2010. https://pslcdatashop.web.cmu.edu/KDDCup/downloads.jsp
Pardos, Z.A., Heffernan, N.T.: Modeling individualization in a Bayesian networks implementation of knowledge tracing. In: De Bra, P., Kobsa, A., Chin, D. (eds.) User Modeling, Adaptation, and Personalization, pp. 255–266. Springer, Berlin (2010)
Pardos, Z. A., & Heffernan, N. T.: KT-IDEM: introducing item difficulty to the knowledge tracing model. In: Konstan, J. A., Conejo, R., Marzo, J. L., & Oliver, N. (eds.) User Modeling, Adaption and Personalization—19th International Conference, UMAP 2011, Girona, Spain. Proceedings (Vol. 6787, pp. 243–254). Springer (2011). https://doi.org/10.1007/978-3-642-22362-4_21
Pardos, Z. A., Trivedi, S., Heffernan, N. T., & Sárközy, G. N.: Clustered knowledge tracing. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia K. (eds.) Intelligent Tutoring Systems—11th International Conference, ITS 2012, Chania, Crete, Greece, June 14–18, 2012. Proceedings (Vol. 7315, pp. 405–410). Springer (2012). https://doi.org/10.1007/978-3-642-30950-2_52

Pardos, Z. A., Bergner, Y., Seaton, D. T., & Pritchard, D. E.: Adapting Bayesian knowledge tracing to a massive open online course in edX. In: S. K. D'Mello, R. A. Calvo, & A. Olney (Eds.), Proceedings of the 6th International Conference on Educational Data Mining, Memphis, Tennessee, USA, pp. 137–144. International Educational Data Mining Society (2013) https://educationaldatamining.org/EDM2013/papers/rn/paper/21.pdf

Pardos and Heffernan (2010b) https://dblp.org/rec/conf/edm/PardosH10.html

Pavlik, P. I., Cen, H., & Koedinger, K. R.: 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, pp. 531–538 (2009)

Pavlik, P., Brawner, K., Olney, A., & Mitrovic, A. (2013). A review of learner models used in intelligent tutoring systems. In: Design Recommendations for Intelligent Tutoring Systems—Learner Modeling: Vol. I pp. 39–68. Army Research Labs, University of Memphis.

Pelánek, R.: Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques. User Model. User-Adap. Inter. **27**(3–5), 313–350 (2017). https://doi.org/10.1007/s11257-017-9193-2

Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L. J., & Sohl-Dickstein, J.: Deep knowledge tracing. Advances in Neural Information Processing Systems, vol. 28. (2015) https://papers.nips.cc/paper/2015/hash/bac9162b47c56fc8a4d2a519803d51b3-Abstract.html

Pu et al. (2019) https://dblp.org/rec/conf/bigdataconf/PuWHC18.html

Qiu, Y., Qi, Y., Lu, H., Pardos, Z. A., & Heffernan, N. T.: Does time matter? modeling the effect of time with bayesian knowledge tracing. EDM (2011)

Ramírez Luelmo, S. I., El Mawas, N., & Heutte, J.:. Machine learning techniques for knowledge tracing: a systematic literature review. In: Proceedings of the 13th International Conference on Computer Supported Education, pp. 60–70. (2021) https://doi.org/10.5220/0010515500600070

Rau, M. A., & Pardos, Z. A.: Adding eye-tracking AOI data to models of representation skills does not improve prediction accuracy. In: Barnes, T., Chi, M., Feng, M. (eds.), Proceedings of the 9th International Conference on Educational Data Mining, EDM 2016, Raleigh, North Carolina, USA, June 29—July 2, 2016 (pp. 622–623). International Educational Data Mining Society (IEDMS). (2016). https://dblp.org/rec/conf/edm/RauP16.html

Ritter et al. (2009) https://dblp.org/rec/conf/edm/RitterHNDMT09.html

Sani, S.M., Bichi, A.B., Ayuba, S.: Artificial intelligence approaches in student modeling: half decade review (2010–2015). Int. J. Comput. Sci. Netw. **5**(5), 746–754 (2016)

Sao Pedro, M., Baker, R. S., & Gobert, J. D.: Incorporating scaffolding and tutor context into Bayesian knowledge tracing to predict inquiry skill acquisition. In: D'Mello, S. K., Calvo, R. A., & Olney, A. (eds.) Proceedings of the 6th International Conference on Educational Data Mining, Memphis, Tennessee, USA, pp. 185–192. International Educational Data Mining Society (2013). https://dblp.org/rec/conf/edm/PedroBG13.html

Sao Pedro, M., Jiang, Y., Paquette, L., Baker, R. S., & Gobert, J. D.: Identifying transfer of inquiry skills across physical science simulations using educational data mining. In: Penuel, W. R., Jurow, A.S.,. O'Connor K. (eds.) Learning and becoming in practice: Proceedings of the 11th International Conference of the Learning Sciences, ICLS 2014, Boulder, Colorado, USA, June 23–27, 2014. International Society of the Learning Sciences. (2014) https://repository.isls.org/handle/1/1116

Schodde, T., Bergmann, K., & Kopp, S.: Adaptive robot language tutoring based on bayesian knowledge tracing and predictive decision-making. In: Mutlu, B., Tscheligi, M., Weiss, A., Young J.E. (eds.) Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, HRI 2017, Vienna, Austria, pp. 128–136. ACM (2017). https://doi.org/10.1145/2909824.3020222

Self, J.A.: Student models in computer-aided instruction. Int. J. Man Mach. Stud. **6**(2), 261–276 (1974)

Skinner, B.F.: The science of learning and the art of teaching. Harv. Educ. Rev. **24**, 86–97 (1954)

Sleeman, D., Brown, J.S.: Introduction: intelligent tutoring systems: an overview. In: Sleeman, D.H., Brown, J.S. (eds.) Intelligent Tutoring Systems, pp. 1–11. Academic Press, Burlington (1982)

Song, Y., Jin, Y., Zheng, X., Han, H., Zhong, Y., & Zhao, X.: PSFK: a student performance prediction scheme for first-encounter knowledge in ITS. In: Zhang, S., Wirsing, M., Zhang, Z., (eds.) Knowledge Science, Engineering and Management—8th International Conference, KSEM 2015, Chongqing, China, Proceedings Vol. 9403, pp. 639–650. Springer. (2015) https://doi.org/10.1007/978-3-319-25159-2_58

Spaulding, S., Gordon, G., & Breazeal, C. (2016). Affect-aware student models for robot tutors. In: Jonker, C.M., Marsella, S., Thangarajah, J., Tuyls K., (eds.), Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems, Singapore, pp. 864–872. ACM. (2016) http://dl.acm.org/citation.cfm?id=2937050

Sun, S., Hu, X., Bu, C., Liu, F., Zhang, Y., Luo, W.: Genetic algorithm for bayesian knowledge tracing: a practical application. In: Tan, Y., Shi, Y., Niu, B. (eds.) Advances in Swarm Intelligence, pp. 282–293. Springer, USA (2022)

Vandewaetere, M., Desmet, P., Clarebout, G.: The contribution of learner characteristics in the development of computer-based adaptive learning environments. Comput. Hum. Behav. **27**(1), 118–130 (2011). https://doi.org/10.1016/j.chb.2010.07.038

Wang, S., Han, Y., Wu, W., Hu, Z.: Modeling student learning outcomes in studying programming language course. Seventh Int. Conf. Inf. Sci. Techno. (ICIST) **2017**, 263–270 (2017). https://doi.org/10.1109/ICIST.2017.7926768

Wang, Y., & Heffernan, N. T.: The student skill model. In: Cerri, S. A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.), Intelligent tutoring systems—11th International Conference, ITS 2012, Chania, Crete, Greece. Proceedings Vol. 7315, pp. 399–404. Springer. (2012) https://doi.org/10.1007/978-3-642-30950-2_51

Wang, Y., & Heffernan, N. T. (2013). Extending knowledge tracing to allow partial credit: using continuous versus binary nodes. In: Lane, H. C., Yacef, K., Mostow, J., & Pavlik, P. I. (eds.), Artificial Intelligence in Education—16th International Conference, AIED 2013, Memphis, TN, USA, July 9–13, 2013. Proceedings Vol. 7926, pp. 181–188. Springer. https://doi.org/10.1007/978-3-642-39112-5_19

Wang, Y., Heffernan, N. T., & Beck, J. E.: Representing student performance with partial credit. In: Baker, R.S., Merceron, A., Pavlik, P. (eds.), Educational data mining 2010, The 3rd international conference on educational data mining, Pittsburgh, PA, USA. Proceedings pp. 335–336 (2010) www.educationaldatamining.org.

Wang, Z., Zhu, J., Li, X., Hu, Z., & Zhang, M.: Structured knowledge tracing models for student assessment on coursera. In: Haywood, J., Aleven, V., Kay, J., Roll, I. (eds.), Proceedings of the Third ACM Conference on Learning @ Scale, L@S 2016, Edinburgh, Scotland, UK, pp. 209–212. ACM. (2016) https://doi.org/10.1145/2876034.2893416

Wenger, E.: Artificial intelligence and tutoring systems. Morgan Kaufmann Publishers Inc, California (1987)

Woolf, B.P.: AI in education. In: Artificial intelligence–encyclopedias, pp. 434–444. Wiley, Hoboken (1992)

Xu, Y., Chang, K., Yuan, Y., Mostow, J.: EEG helps knowledge tracing ! https://www.semanticscholar.org/paper/EEG-Helps-Knowledge-Tracing-!-Xu-Chang/645cb28a5d3802c14305ae239a27dd3506ec71f9

Xu, Y., Mostow, J.: Using item response theory to refine knowledge tracing. In: D'Mello, S. K., Calvo, R. A., Olney, A. (eds.), Proceedings of the 6th International Conference on Educational Data Mining, Memphis, Tennessee, USA, July 6–9, 2013 pp. 356–357. International Educational Data Mining Society. (2013). https://dblp.org/rec/conf/edm/XuM13.html

Xu, Y., Johnson, M. J., & Pardos, Z. A.: Scaling cognitive modeling to massive open environments. In: Proceedings of the Workshop on Machine Learning for Education at the 32nd International Conference on Machine Learn- Ing (ICML) (2015)

Yudelson, M.: (2016) https://dblp.org/rec/conf/edm/Yudelson16a.html

Yudelson, M., Koedinger, K.R., Gordon, G.J.: Individualized Bayesian knowledge tracing models. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) Artificial Intelligence in Education, pp. 171–180. Springer, Berlin Heidelberg (2013)

Yudelson, M., Medvedeva, O., Crowley, R.S.: A multifactor approach to student model evaluation. User Model. User-Adap. Inter. **18**(4), 349–382 (2008). https://doi.org/10.1007/s11257-007-9046-5

Yudelson, M.: Individualization of Bayesian knowledge tracing through elo-infusion. In: Roll, I., McNamara, D. S., Sosnovsky, S. A., Luckin, R., Dimitrova, V. (eds.) Artificial Intelligence in Education—22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part II Vol. 12749, pp. 412–416. Springer (2021). https://doi.org/10.1007/978-3-030-78270-2_73

Zafar, A., & Ahmad, N.: An overview of Student Modeling Approaches under Uncertain Conditions. Int. J. Inf. Technol. Manage. *4*(1) (2013)

Zhang, K., Yao, Y.: A three learning states Bayesian knowledge tracing model. Knowl. Based Syst. **148**, 189–201 (2018). https://doi.org/10.1016/j.knosys.2018.03.001

Zhou, X., Wu, W., Han, Y.: Modeling multiple subskills by extending knowledge tracing model using logistic regression. IEEE Int. Conf. Big Data (big Data) **2017**, 3994–4003 (2017). https://doi.org/10.1109/BigData.2017.8258413

Zhu, J., Zang, Y., Qiu, H., Zhou, T.: Integrating temporal information into knowledge tracing: a temporal difference approach. IEEE Access **6**, 27302–27312 (2018). https://doi.org/10.1109/ACCESS.2018.2833874

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.