

WHAT IS BAYESIAN KNOWLEDGE TRACING?

YOUNG CHO, GRACE MAZZARELLA, KELVIN TEJEDA, TONGYU ZHOU, IRIS HOWLEY (iris@cs.williams.edu)

LEARNING ANALYTICS SYSTEMS USE BKT TO PREDICT WHAT SKILLS STUDENTS MASTERED

<http://www.cs.williams.edu/~iris/res/bkt/>



A BKT Road Trip

Bayesian Knowledge Tracing (BKT) is an algorithm used in educational systems to predict the likelihood that a student has learned a skill. As the student answers questions, BKT uses different parameters to calculate the chance. This journey tracks understanding as much like a road to the destination as the driver's choice. The driver doesn't know when they're going to get there, but each correct turn they take increases the chance that they'll arrive and every wrong turn increases the time it takes to get there.

At the beginning of the trip, the driver is parked and ready to go. Along with their skill, experience and knowledge they've gathered over the years.

The **Prior Knowledge** is based on BKT. The more the student knows at the start, the closer they are to their destination.

In other words, there is a higher probability that the student has already learned the skill.

The driver is focused with their first choice left or right?

This part of the route is familiar and they choose correctly. Left.

In BKT, correct answers increase the chance that the student has learned the skill.

The driver has been driving for a while now, choosing the correct at every fork. Each **Correct Choice** takes them closer to their destination.

They choose **RIGHT** instead of the correct **LEFT** turn that they meant to take. This error is called a **Slip** and takes them off course from their destination.

When BKT predicts that the student has a **Slip** error, it assumes that the student does know the right answer, but the overall will introduce uncertainty into the overall knowledge prediction.

The trip will take a bit longer now.

Now, they're at an intersection. They have three choices: straight, left or right. The driver doesn't remember this part of the journey.

They guess, continuing straight. The journey will take even longer now.

In BKT when a student answers incorrectly this may be either a **Slip** or a **Guess**. A **Slip** predicts that the student knows the information, but made a mistake answering the question. A **Guess** predicts that the student does not know the information.

They try to turn right again. The destination is approaching and a fork in the road shows how to get there after all.

In BKT many correct answers increase the odds that the student knows the skill, until eventually a certain level is reached, and the algorithm determines that the skill has been learned.

They've arrived!

The trip took a little longer than expected, with the occasional **Slip** and the **Guess** the driver made, but in the end, they did arrive that they know how to get to their destination. Their **Prior Knowledge** started, and every **Correct Choice** they made helped them overcome their small mistakes.

WHY BAYESIAN KNOWLEDGE TRACING?

- Maintains an estimate of the probability that students have learned a particular set of skills in Technologically Enhanced Learning Environments
- Used in the Open Learning Initiative, Open Analytics Research Service, etc. within Learning Analytics systems
- Instructors rely on BKT to assist in making decisions in the classroom
 - Concepts to review
 - Students to follow up with
 - Exam question identification

A 2-node dynamic Bayesian network

$$P(L_1) = P(L_0) \quad (1)$$

$$P(L_{n-1}|obs_n = corr) = \frac{P(L_{n-1}) * (1 - P(S))}{P(L_{n-1}) * (1 - P(S)) + (1 - P(L_{n-1})) * P(G)} \quad (2)$$

$$P(L_{n-1}|obs_n = incorr) = \frac{P(L_{n-1}) * P(S)}{P(L_{n-1}) * P(S) + (1 - P(L_{n-1})) * (1 - P(G))} \quad (3)$$

$$P(L_n|obs_n) = P(L_{n-1}|obs_n) + ((1 - P(L_{n-1}|obs_n)) * P(T)) \quad (4)$$

$P(L_0)$: Probability student already knew the skill
 $P(T)$: Probability student learned after learning opportunity
 $P(G)$: Probability student guessed correctly
 $P(S)$: Probability student made a mistake on known skill
 If $P(L_n) > 0.95 \rightarrow$ Skill is mastered

FUTURE WORK

- Interactivity of Explainables \rightarrow Understanding the Algorithm
- Understanding the Algorithm \rightarrow Trust, Fairness, Model Interrogation, Decision-making, ...

DESIGN PRINCIPLES FROM THE LEARNING SCIENCES

- Learning by doing:** Learning increases with interactive activities vs. watching videos
- Backward Design:** Identify learning goals first, then assessments *then* learning activities
- Cognitive Task Analysis:** Systematically identify hierarchical skills & concepts to learn
- Zone of Proximal Development:** 'Goldilocks' of learning challenging concepts

DOES UNDERSTANDING THE ALGORITHM MATTER, IF THE USER CAN'T CHANGE THE SYSTEM?

EXPLAINABLES

Post-hoc Interpretability of Artificial Intelligence Algorithms
Approaches from Lipton (2017)

Generated Text Explanations

SAMPLE SCREENSHOTS

Visual Cooking Narrative implemented in ren.py

Interactive Alchemy implemented in HTML/CSS:

DESIGN PROCESS

Brainstorm \leftrightarrow Paper Prototype \leftrightarrow Usability Testing \leftrightarrow Implement \leftrightarrow Usability Testing

Themes: Consider individual differences, Refine the level of detail, Usability design principles

- "Am I supposed to be remembering all these [parameters]? Because that ain't gonna happen."
- "I was able to understand it more because it was more in depth."
- Yeah, now I think I'd believe that [BKT works]."

Affinity Diagram of User Comments

REFERENCES

OARS: exploring instructor analytics for online learning by Bassen, Howley, Fast, Mitchell, & Thille (2018).
The myths of model interpretability by Lipton (2016).

Knowledge tracing: Modeling the acquisition of procedural knowledge tracing by Corbett, & Anderson (1995).
Individualized bayesian knowledge tracing models by Yudelson, Koedinger, & Gordon (2013).
More accurate student modeling through contextual estimation of slip and guess probabilities in bayesian knowledge tracing by Baker, Corbett, & Alevan (2008).