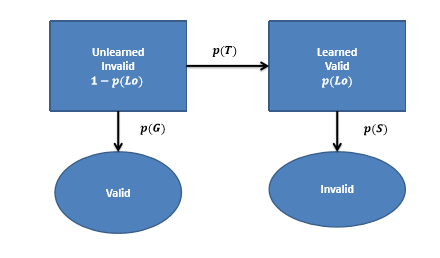
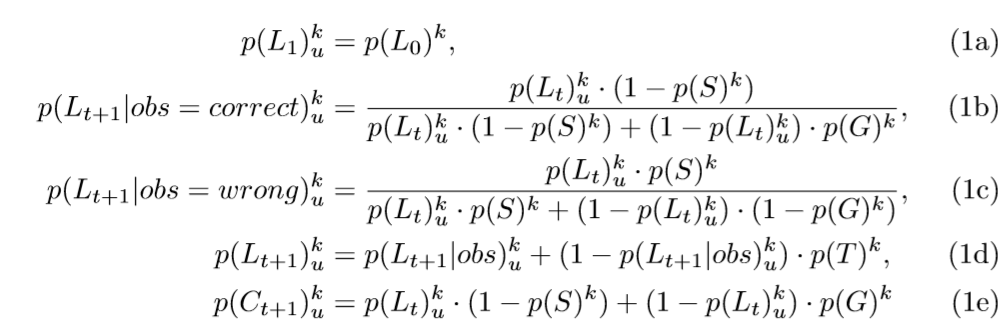
Bayesian Knowledge Tracing (BKT)

Bayesian Knowledge Tracing (BKT) is a user modelling method extensively used in the area of Intelligent Tutoring Systems. In the standard BKT implementation, there are only skill-speciﬁc parameters. BKT assumes that student knowledge is represented as a set of binary variables – one per skill (the skill is either mastered by the student or not). Observations in BKT are also binary: a student gets a problem either right or wrong.

The BKT model is built as a dynamic Bayesian network of two states, as shown in figure, which takes observations about student performance in solving problems where a skill to be learned is involved (correct or incorrect answer). The model assumes that at any given opportunity to demonstrate a skill, a student knows the skill (learned) or do not know the skill (unlearned), and can either take a right or wrong answer. A student who has not learned a skill usually gives a wrong answer; however, there is a certain probability **P(G)** that the student gives a correct answer (this is the guess parameter). Similarly, a student who has not learned a skill usually gives a correct answer; in this case there is a certain probability **P(S)** that the student is wrong (slip). The ﬁrst time the model is used, each student has an initial probability **P(Lo)** to learn a skill, and every time where the student has an opportunity to practice skill there is a certain probability **P(T)** to make a transition in the state, from unlearned to learned.



The estimate of the level of the student knowledge of a skill is continually updated each time the student gives an answer. right or wrong, to a problem involving the skill by following equations:



The initial probability of student **u** mastering skill **k** is set to the **p-init** parameter for that skill Equation (1a). Depending on whether the student **u** applied skill **k** correctly or incorrectly, the conditional probability is computed either using Equation (1b) or Equation (1c). The conditional probability is used to update the probability of skill mastery according to Equation (1d). To compute the probability of student u applying the skill k correctly on an upcoming practice opportunity one uses Equation (1e).

To apply above mentioned equations, we first need the four types of model parameters used in BKT. For this purpose, approach of **Empirical Probabilities(EP)** is used. EP is a simple, two-phase method for fitting Knowledge Tracing models. The first phase involves annotating performance data with knowledge using a heuristic, and the second phase takes these knowledge labels along with student performance and empirically computes the four KT parameters.

**Knowledge Heuristic**

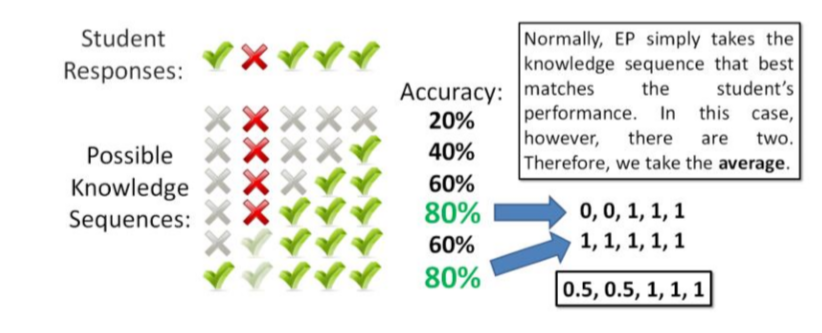
Normally, knowledge is unobservable when training a KT model. However, a heuristic makes it observable to allow the four KT parameters to be computed empirically. The way this works is that it takes a student’s sequence of correct and incorrect responses for a given skill (performance sequence) and estimates the point at which the student learned the skill. The student is considered to be in the “unknown state” on all questions before that point, and in the “known state” on all subsequent questions, since we assume that no forgetting takes place. We refer to the resulting sequence of unknown and known states as the knowledge sequence for that student on that skill.

To determine when the student learned the skill, the heuristic tries all possible knowledge sequences that are the same length as the student’s performance sequence, and selects the one that matches up with the performance sequence the best. We consider a known state to match up with a correct response, and an unknown state to match up with an incorrect response. For example, if a student’s performance sequence is (correct, incorrect, incorrect, correct, correct) then the knowledge sequence that matches up with it the best is (unknown, unknown, unknown, known, known) , since it matches up with four of the five performances correctly, more than any other possible knowledge sequence does.

Steps included in this approach is as follows:-

1. **Annotating Knowledge**:

The first step in EP is to annotate performance data for each student within each skill with an estimate of when the student learned the skill. We assume there are only two knowledge states: known (1) and unknown (0), and do not allow for forgetting (a known state can never be followed by an unknown state). For this we choose the knowledge sequence that best matches their performance.



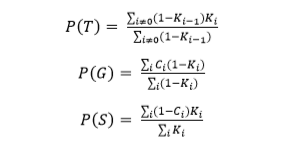
1. **Computing the Probabilities**

Using the knowledge estimates, we can compute each of the four BKT parameters for each skill empirically from the data.

The first of these parameters is **P(L0)**, the probability that the student knew the skill before interacting with the system. We can empirically estimate this by taking the average value of student knowledge on the first practice opportunity:



Using **K(i)** and **C(i)** as knowledge and correctness at problem **i,** respectively, the following equations are used to compute the other three BKT parameters:



Now that we have all the model parameters, we can apply equations (1a) to (1e) and can predict if a student has gained mastery over a skill or not.