

# Learning to Teach Using Reinforcement Learning

Susanna Weinberger & Michael Raffelsberger



## Project Overview

Data Science offers the possibility to improve education through **personalized tutoring systems**.



## Goal

Build a tutoring system to **infer students' knowledge** and learn to **provide good teaching recommendations**.

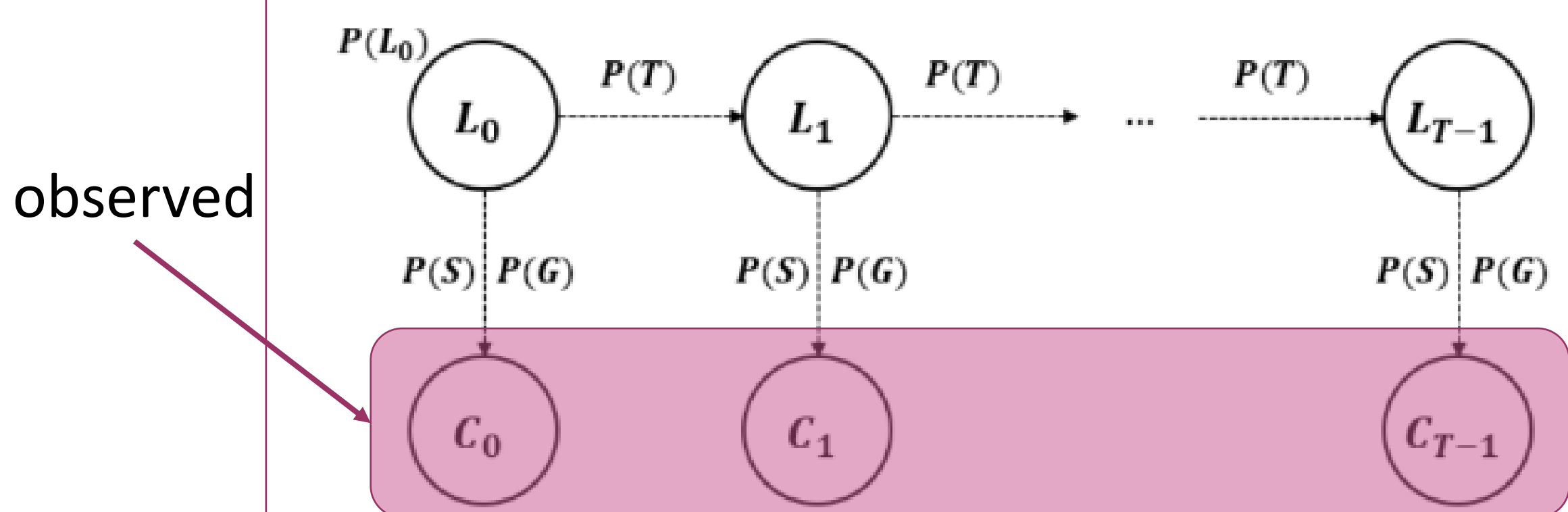
## Knowledge Tracing

### Parameters

$P(L_0)$  ... prior probability  
 $P(T)$  ... transition (=learning) probability  
 $P(G)$  ... guess probability  
 $P(S)$  ... slip probability

### Variables

$L_t \in \{0,1\}$  ... learning state (latent)  
 $C_t \in \{0,1\}$  ... correct flag (observed)



Use a **Hidden Markov Model (HMM)** to estimate the parameters from observed exercised sequences from students.

### How to use the estimated parameters?

$$P(L_{t+1} = 1) = P(L_t = 1|C_t) + (1 - P(L_t = 1|C_t)) \cdot P(T)$$

$$P(L_t = 1|C_t) = \frac{P(C_t|L_t = 1) \cdot P(L_t = 1)}{P(C_t)}$$

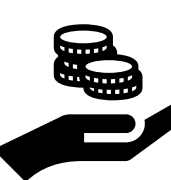
We extended the standard BKT model above for a single skill to a setting where we have **blocks of dependent skills**, e.g.:

- *Block 1: deriving functions, square roots, slope of a line*
- *Block 2: expected values, conditional probability, bayes' theorem*

## Reinforcement Learning



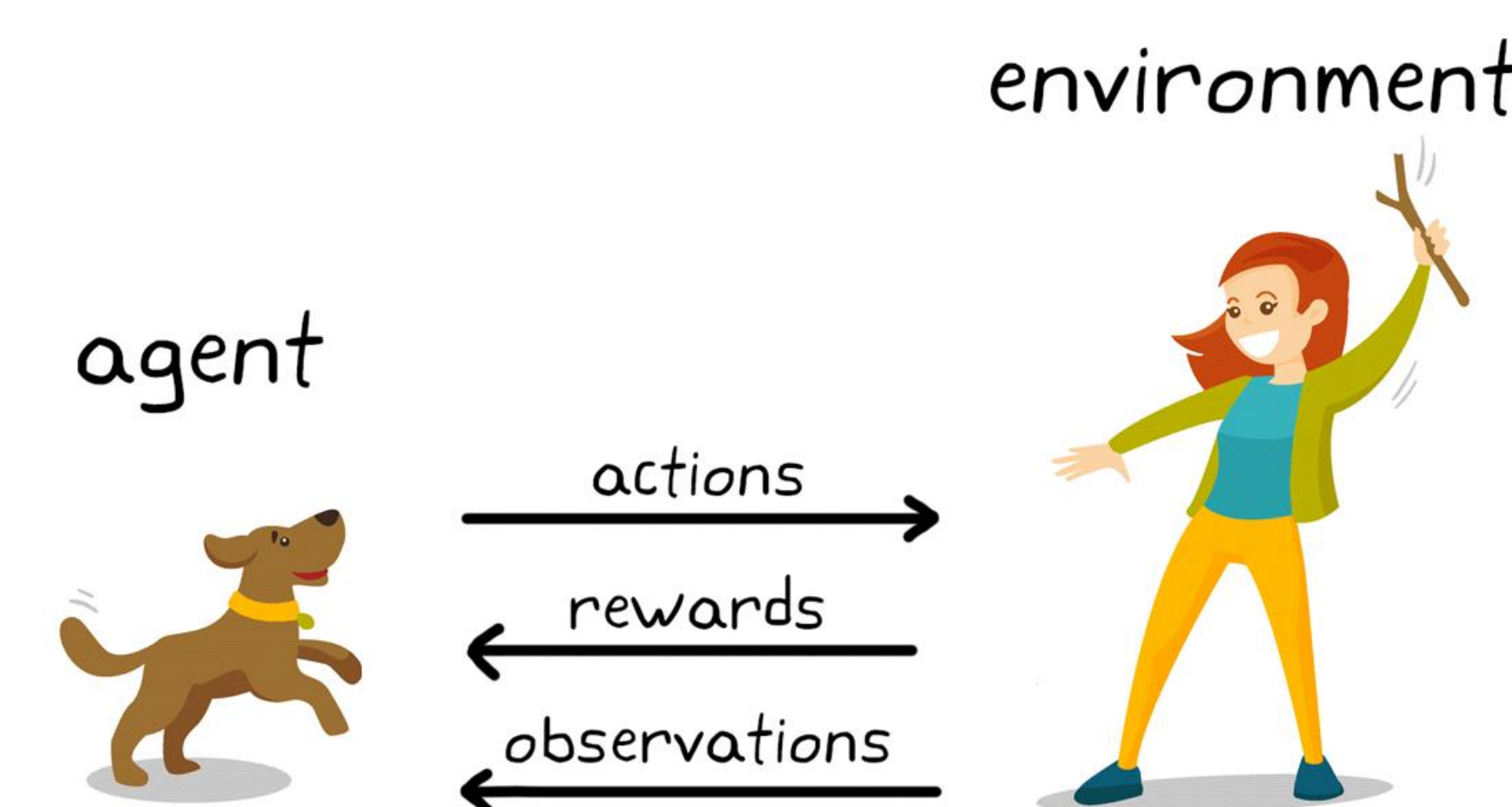
**Action:** Which exercise / skills should be taken.



**Reward:** Multiple possibilities, but dependent on learned skills.



**Observation:** (Probabilistic) learning state of the student.

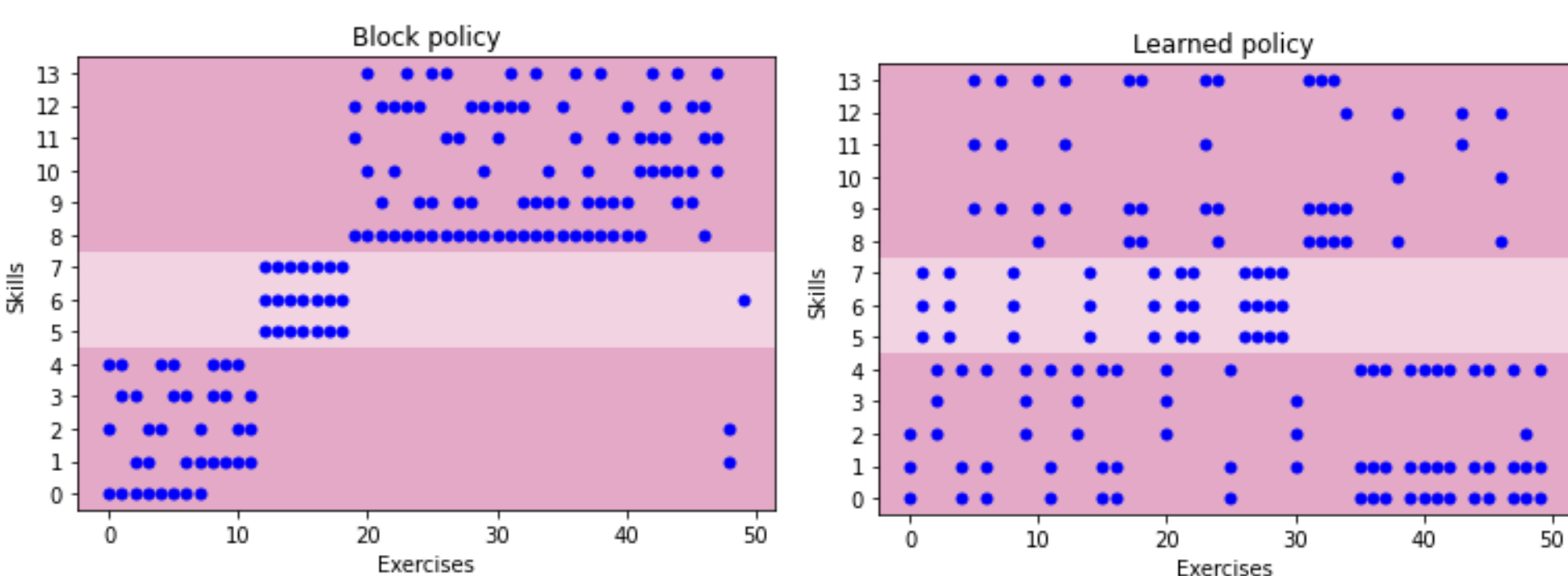


We experimented with different configurations to improve our performance. Student received 50 exercises from 14 learnable skills.

	Agent 1	Agent 2	Agent 3
Reward Agent	369.37	265.45	11.38
Reward Best Basel.	320.09	309.64	11.22
Skills Learned Agent	10.32	6.69	11.38
Skills Learned Best Basel.	10.18	9.83	11.22

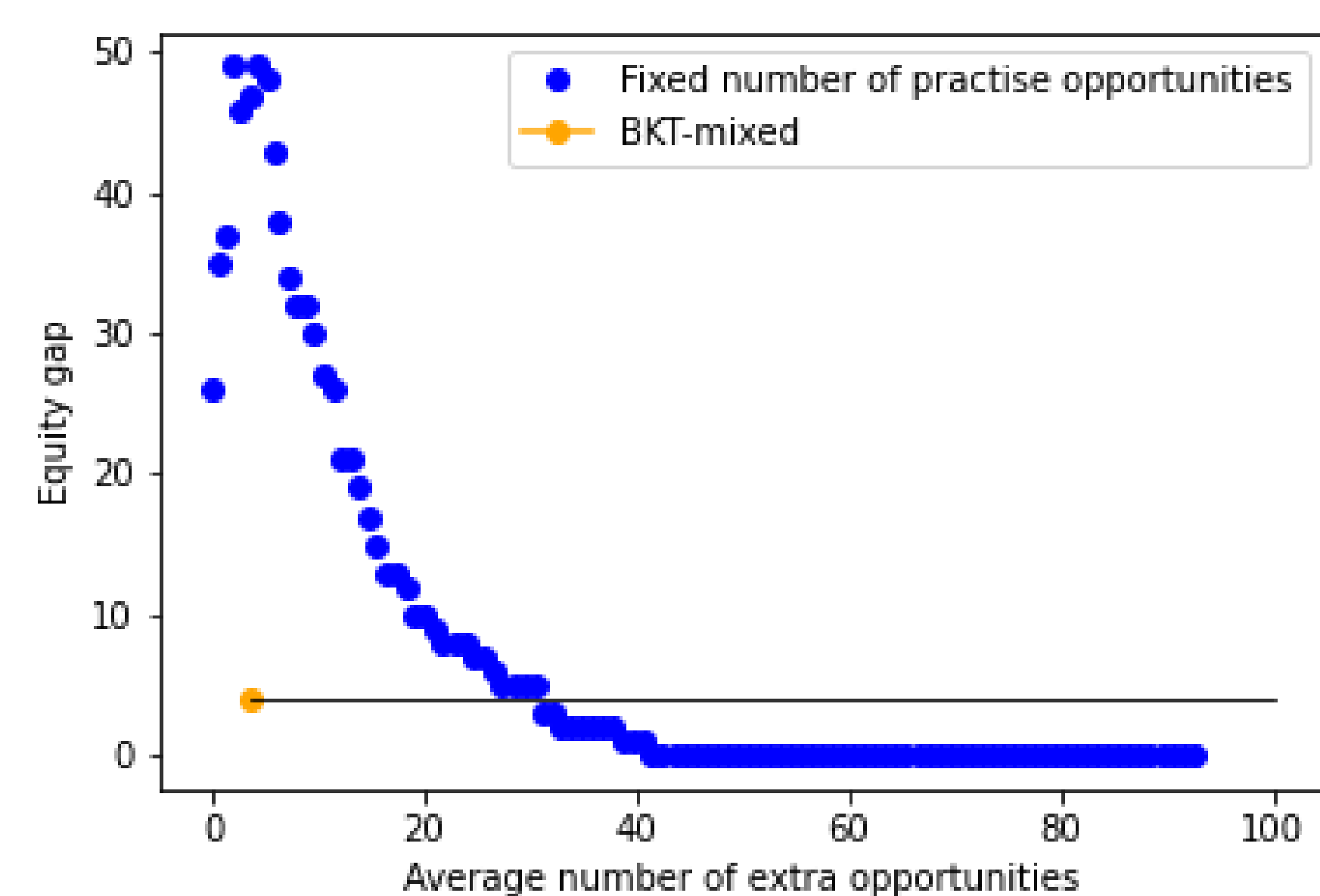
## Curricula Experiments

It is Interesting to look at are the actual actions the agent takes (so the curriculum each student receives).



## Fairness

Students are no homogenous group, so learn at a different speed – we would like to provide everyone with a good learning success, while optimizing the time spent for learning (number of exercises).



We looked at two student groups – slow and fast learners and compared how well they learned with BKT and when giving them a fixed amount of exercises.