Learning Student and Content Embeddings for Personalized Lesson Sequence Recommendation

Siddharth Reddy, Igor Labutov, Thorsten Joachims

{sgr45, iil4}@cornell.edu, tj@cs.cornell.edu

Key Contributions

- Demonstrate the ability of an embedding model to successfully predict assessment results
- Introduce an offline methodology as a proxy for assessing the ability of a model to recommend personalized lesson sequences

Model Representation

- **Student** = a set of d latent skill levels $\vec{s} \in \mathbb{R}_+^d$ that vary over time
- **Lesson module** = a vector of skill gains $\vec{\ell} \in \mathbb{R}_+^d$ and a set of prerequisite skill requirements $\vec{q} \in \mathbb{R}_+^d$
- **Assessment module** = a set of skill requirements $\vec{a} \in \mathbb{R}_+^d$
- A student can be tested on an assessment module, which has a **pass-fail result** $R \in \{0, 1\}$. The likelihood of passing should be high when a student has skill levels that exceed the assessment requirements, and vice-versa.
- A student can complete lesson modules to learn over time, though the **skill gains** $\vec{\ell}$ from a lesson module are modulated by **prerequisite knowledge** $\vec{q}$

Model Dynamics

Assessment Results

For student $\vec{s}_t$, assessment $\vec{a}$, and result R,

$$R \sim \text{Bernoulli}(\phi(\Delta(\vec{s}_t, \vec{a})))$$

where ϕ is the logistic function and

$$\Delta(\vec{s}_t, \vec{a}) = \frac{\vec{s}_t \cdot \vec{a}}{||\vec{a}||} - ||\vec{a}|| + \gamma_s + \gamma_a$$

Student Learning from Lessons

For student $\vec{s}$ who worked on a lesson with skill gains $\vec{\ell}$ and prerequisites $\vec{q}$ at time $t+1$, the updated student state is

$$\vec{s}_{t+1} \sim \mathcal{N}\left(\vec{s}_t + \vec{\ell} \cdot \phi(\Delta(\vec{s}_t, \vec{q})), \Sigma\right)$$

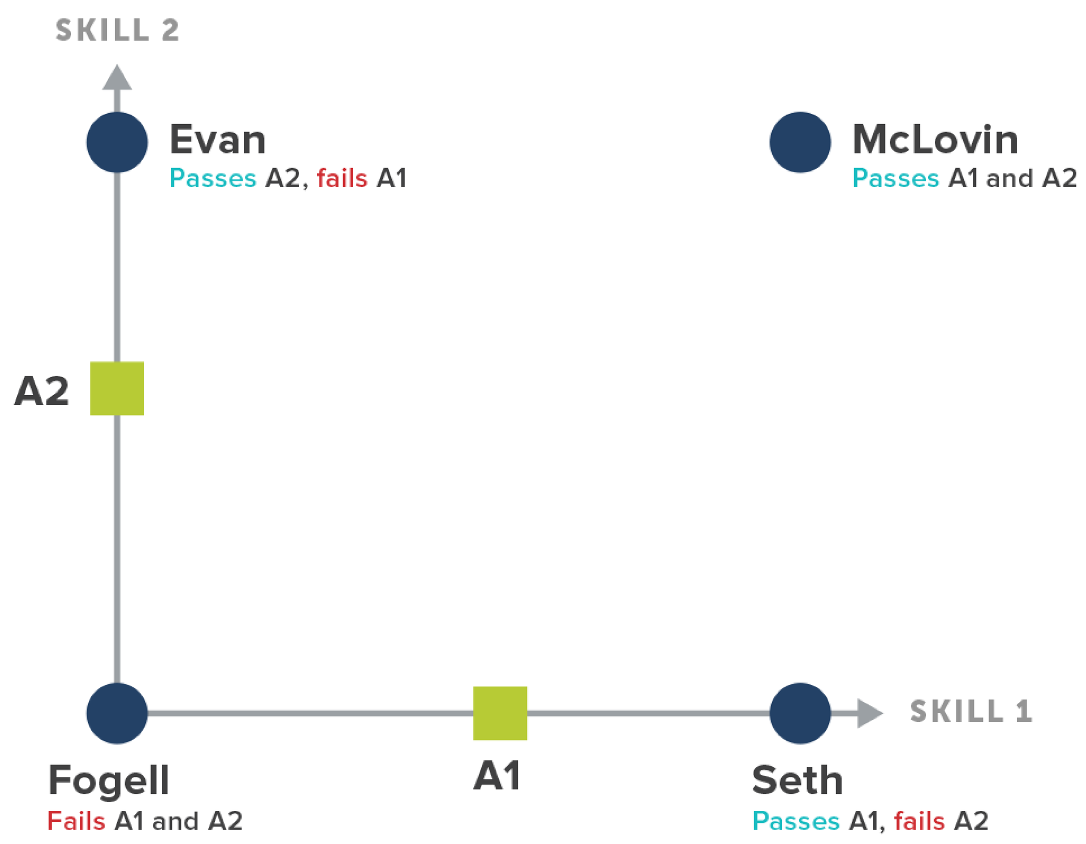
where the covariance matrix is $\Sigma = I_d \sigma^2$

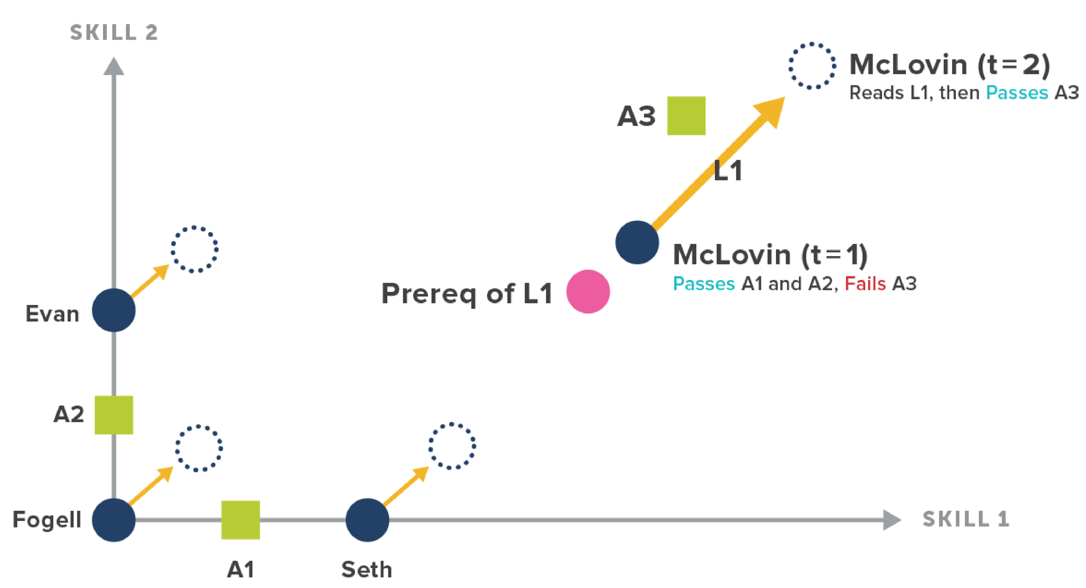
Parameter Estimation

We compute MAP estimates of model parameters Θ by maximizing the following objective function:

$$L(\Theta) = \sum_{\mathcal{A}} \log\left(\mathbb{P}[R \mid \vec{s}_t, \vec{a}, \gamma_s, \gamma_a]\right) + \sum_{\mathcal{L}} \log\left(\mathbb{P}[\vec{s}_{t+1} \mid \vec{s}_t, \vec{\ell}, \vec{q}]\right) - \beta \cdot \lambda(\Theta)$$

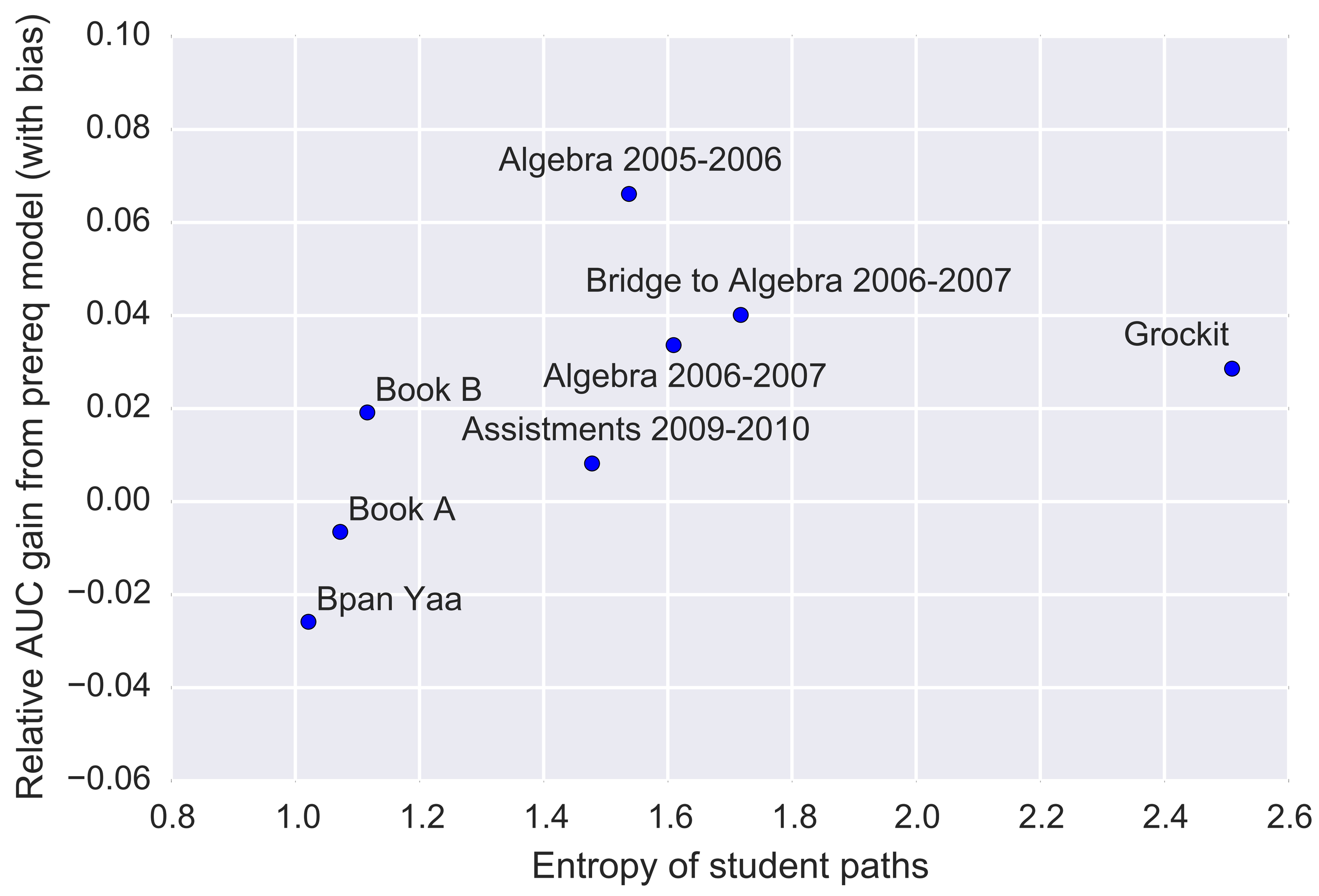
Examples

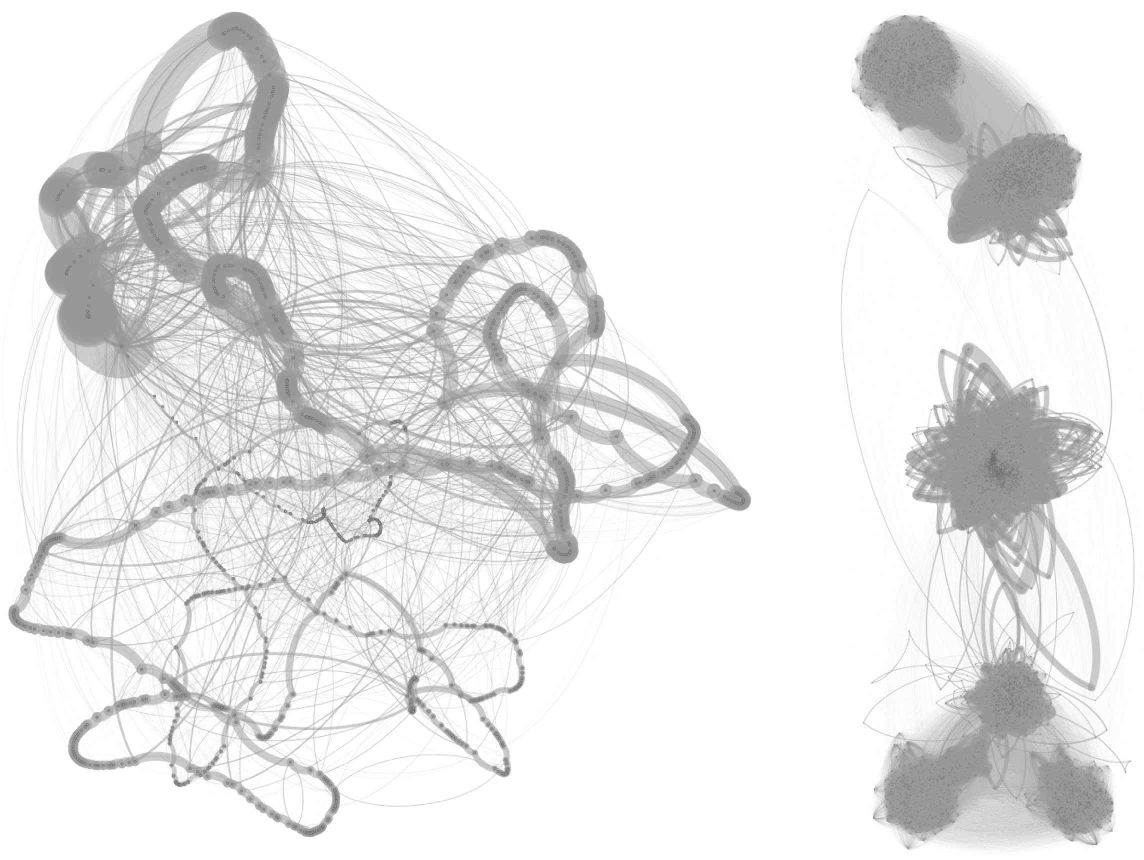


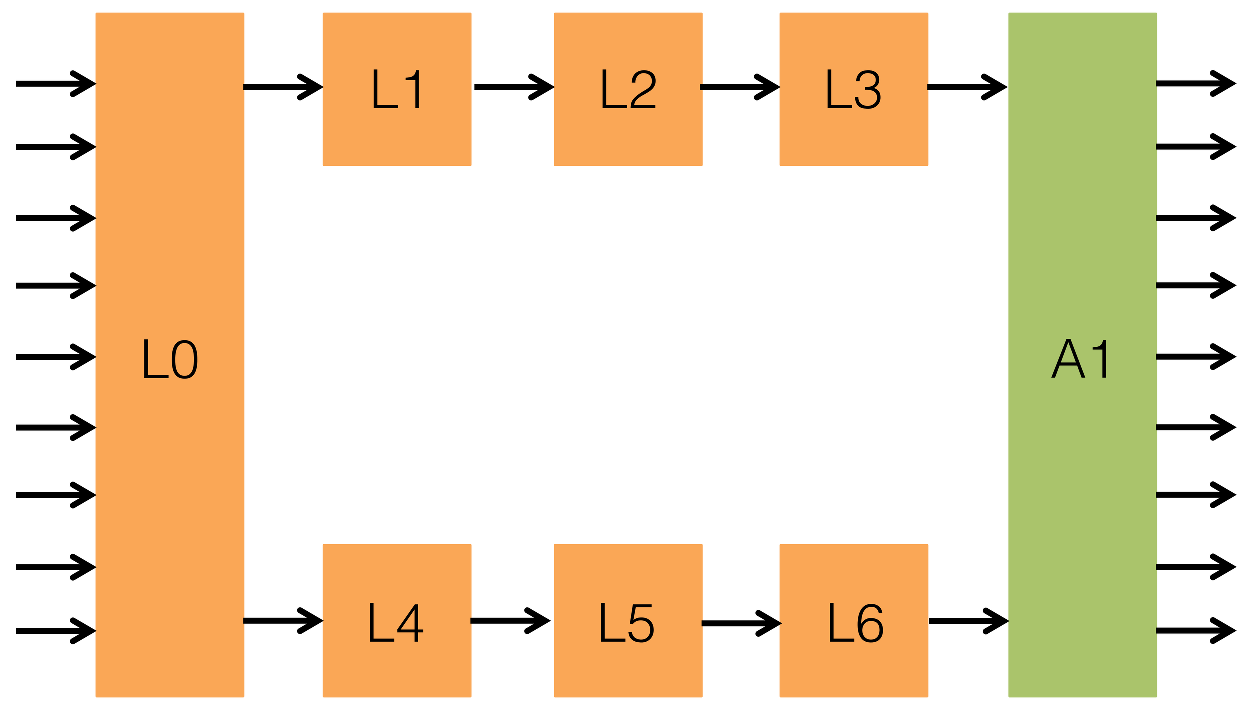


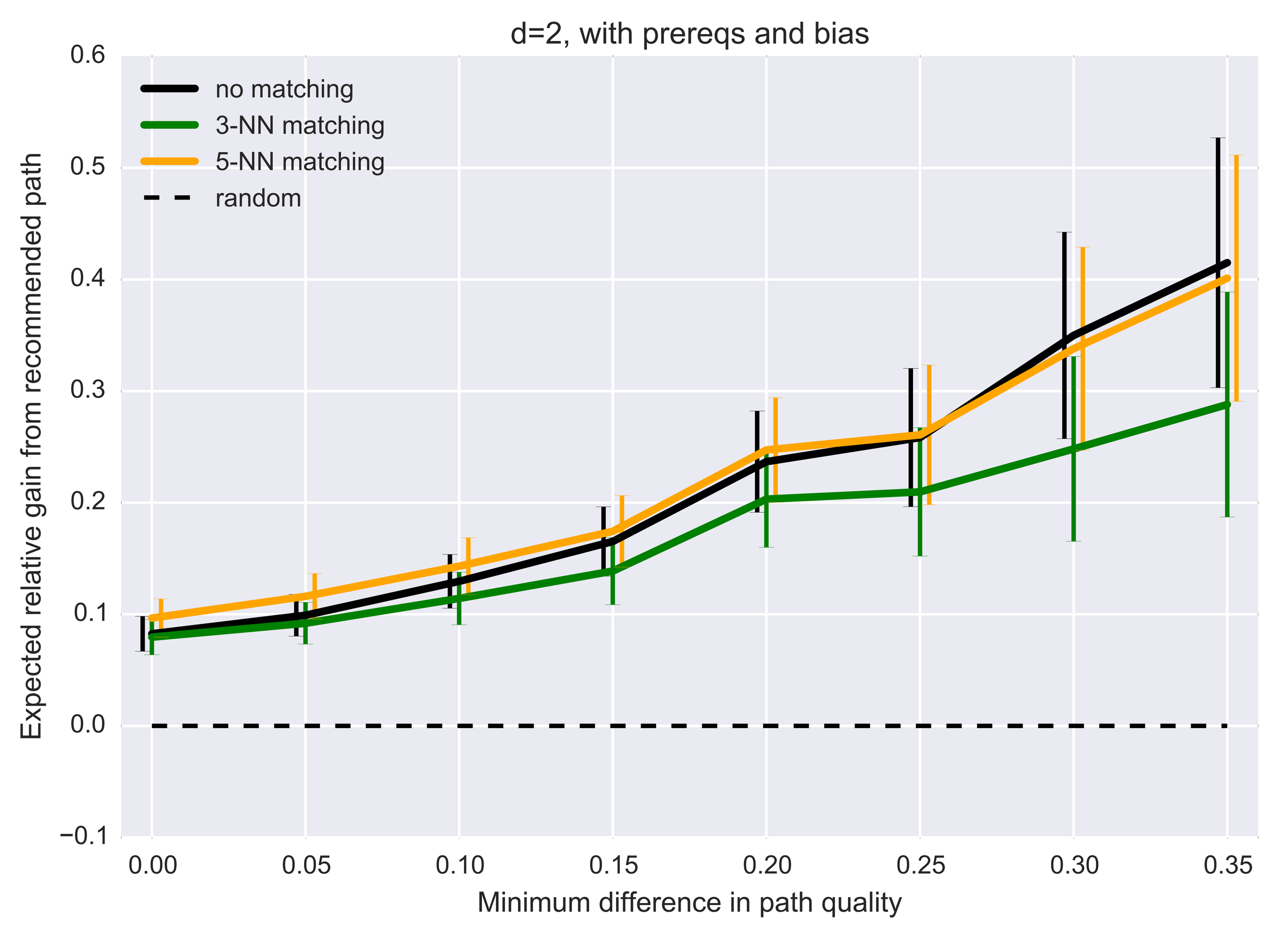
Experiments on Online Course Data

	Model			Book A		Book B	
	$\vec{\ell}$	$\vec{q}$	γ	Test	Validation	Test	Validation
1	N	N	N	0.673	0.614 ± 0.015	0.614	0.644 ± 0.015
2	N	N	Y	0.818	0.753 ± 0.020	0.788	0.821 ± 0.021
3	Y	N	N	0.692	0.624 ± 0.019	0.630	0.662 ± 0.023
4	Y	N	Y	0.798	0.761 ± 0.016	0.775	0.808 ± 0.020
5	Y	Y	N	0.724	0.625 ± 0.021	0.629	0.643 ± 0.018
6	Y	Y	Y	0.811	0.756 ± 0.018	0.785	0.823 ± 0.021
7	1PL IRT			0.812	0.761 ± 0.016	0.778	0.812 ± 0.019
8	2PL IRT			0.780	0.708 ± 0.011	0.686	0.690 ± 0.022
9	2D MIRT			0.817	0.732 ± 0.012	0.776	0.796 ± 0.018









Full paper and code available at `http://siddharth.io/lentil`