Carnegie Mellon University

HCI Tutor Research

Progress Report

Elaine Xu, Smeet Poladia, Zhou Lu

Master of Statistical Practice students, Department of Statistics and Data Science

Introduction

Zhou Lu

Statistics & Machine Learning
Economics

Smeet Poladia

Statistics and Machine Learning

Elaine Xu

Statistics & Machine Learning
Cognitive Science

Carnegie Mellon University

Background

- Client: Dr. Vincent Aleven, HCI Institute, CMU
- It has been established that students learn faster and better using Intelligent Tutoring Systems (ITSs). In these systems, data about students, assessments, learning, etc is logged.
- The project aims to investigate whether prerequisite relations among math topics can be detected in log data.

Client Info

Vincent Aleven

- Background: CS, Intelligent Systems
- Co-founder of Carnegie Learning & MathTutor
- AI-based tutoring software

Carnegie Mellon University

Our Understanding

Scope:
Develop a method for finding prerequisite relations among math topics/skills/units, given log data from students doing tutor work on these topics/skills/units.

Challenges:

- Test whether topic/skill/unit A is prerequisite for B
  - Treat “Strong Correlations” as evidence of a prerequisite relation
- What metrics of learning and performance?
- At what level of granularity?
  - Start small, start simple

Approach

High level Project Plan	
Project Definition and Value Proposition	February
Data Collection, Data Pre-processing, EDA	February, March
Understanding Correlations, Building a Statistical Model, Developing Metric	March, April
Final Deliverables & Project Debrief	May

MathTutor

Find some common multiples of 8 and 14. Then, find the least common multiple of 8 and 14.

*Enter the multiples of each number, in order. Then click the checkbox to the right of a multiple if the number is a common multiple.

Multiples of 8 in order:	Multiples of 14 in order:
8	14
16	28
24	42
32	56
40	70
48	84
56	98
64	112
72	
80	

The least common multiple of 8 and 14 is:

Hint

Enter Multiples
Identify Common Multiples
Enter Least Common Multiple

https://mathtutor.andrew.cmu.edu/

Carnegie Mellon University

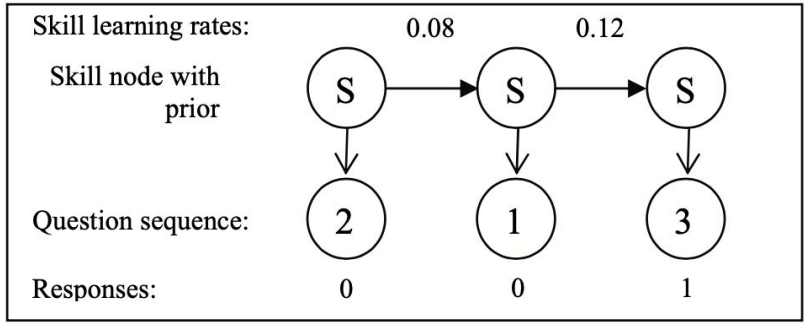
DataShop

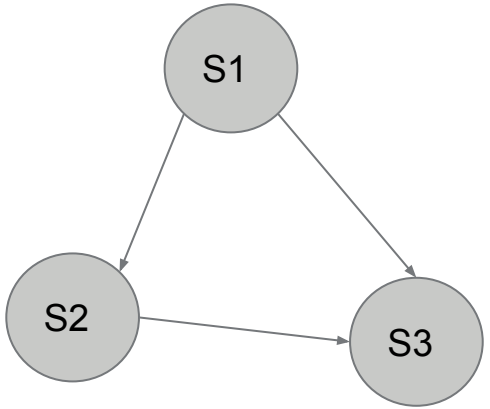
- DataShop
  - Interaction between students and educational software

#	Student	Problem	Step	Attempt #	Student Input	Evaluation	Knowledge component
1	S01	MAKING-CANS	(SQUARE-BASE Q1)	1	8	CORRECT	Enter-Given
2	S01	MAKING-CANS	(SCRAP-METAL-AREA Q1)	1	32	INCORRECT	
3	S01	MAKING-CANS	(SCRAP-METAL-AREA Q1)	2	4	INCORRECT	
4	S01	MAKING-CANS	(SQUARE-AREA Q1)	1	64	CORRECT	Square-Area
5	S01	MAKING-CANS	(POG-AREA Q1)	1	50.24	CORRECT	Circle-Area
6	S01	MAKING-CANS	(SCRAP-METAL-AREA Q1)	3	13.76	CORRECT	Compose-Areas
7	S01	MAKING-CANS	(POG-RADIUS Q2)	1	8	CORRECT	Enter-Given
8	S01	MAKING-CANS	(SQUARE-BASE Q2)	1	16	CORRECT	Enter-Given
9	S01	MAKING-CANS	(SQUARE-AREA Q2)	1	256	CORRECT	Square-Area

Koedinger, K. R., Baker, R. S. J. d., Cunningham, K., Skogsholm, A., Leber, B., & Stamper, J. (2010). A data repository for the EDM community: The PSLC datashop. In S. Ventura, C. Romero, M. Pechenizkiy, & R. S. J. d. Baker (Eds.), Handbook of educational data mining (pp. 43–55). Boca Raton, FL: CRC Press.

Literature Review





Left: Significance of item order in randomized problem sets

Right: Bayesian Network

Z. A. Pardos and N. T. Heffernan. Determining the significance of item order in randomized problem sets. 2009.

Carnegie Mellon University

Thank you for your time!