Distributed Representation of Misconceptions

Paper link: http://tiny.cc/**icls_distributed**

Zachary A. Pardos
University of California at Berkeley

Scott Farrar
Khan Academy

John Kolb, Gao Xian Peh, Jong Ha Lee
University of California at Berkeley





Computational Approaches to Human Learning (CAHL) research lab

Do these answers share a misconception?

(difficult task)

Find the missing fraction.

$$\frac{5}{6} - \boxed{4/3} = \frac{1}{2}$$

problem B2

Find the missing fraction.

$$\frac{5}{2} - \boxed{17/6} = \frac{1}{3}$$

problem B1

Yes, they both involve the procedure of adding the left and right side of the Eq.

Main Objectives of the Research

1. Leverage big data to algorithmically cluster wrong answers, like these, together by misconception

2. Design interventions that adhere to guidance from education theory which can scale if the big data approach is successful

Brief Synopsis of Literatures on Misconceptions

Education theory / learning science

Piaget, Smith, diSessa, Roschelle, Franke, Sfard, Cobb…

- Stress treatment of misconceptions as partial understanding not to be confronted

"*Persistent misconceptions, if studied in an evenhanded way, can be seen as novices' efforts to extend their existing useful conceptions to instructional contexts in which they turn out to be inadequate*" - Smith et al. (1994)

Structure, connectionist (continuous conceptual relations)

Computational cognitive science

Anderson, Corbett, VanLehn, Koedinger, Brown, Sleeman…

- Treat misconceptions as "*buggy rules,*" representing incorrect variations of the correct reasoning processes in the ideal model within an Intelligent Tutoring System (ITS)

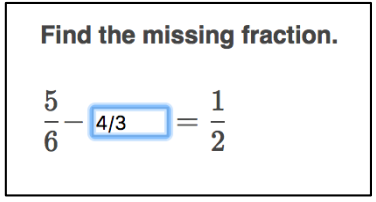
Rule-based, symbolic (discrete tagging)

Crossover between fields:

The *big data* algorithm we adopt to represent wrong answers comes from the *computational sciences* but lends itself well to ideas from *education theory*

Algorithmic Approach to Answer Clustering

Task: learn representation of the answer of “4/3” to problem B2 then cluster similar representations



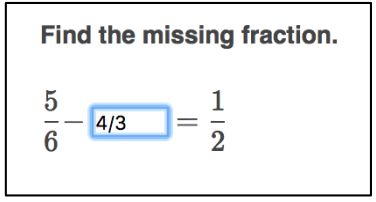
problem B2

Student sequence of answers to problems in Fraction Subtraction exercise:

problem_A1 (35.5), problem_C2 (17/3), problem_B2 (4/3), problem_C1 (-55), problem_D1 (9/2)

Algorithmic Approach to Answer Clustering

Task: learn representation of the answer of “4/3” to problem B2 then cluster similar representations



problem B2

Student sequence of answers to problems in Fraction Subtraction exercise:

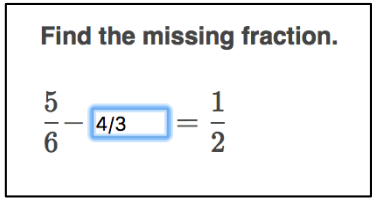
problem_A1 (35.5) problem_C2 (17/3) problem_C1 (-55) problem_D1 (9/2)

Intuition: problem answers in the context of problem_B2 serve as a *signature* of the misconception

problem_B2 (4/3)

Algorithmic Approach to Answer Clustering

Task: learn representation of the answer of “4/3” to problem B2 then cluster similar representations

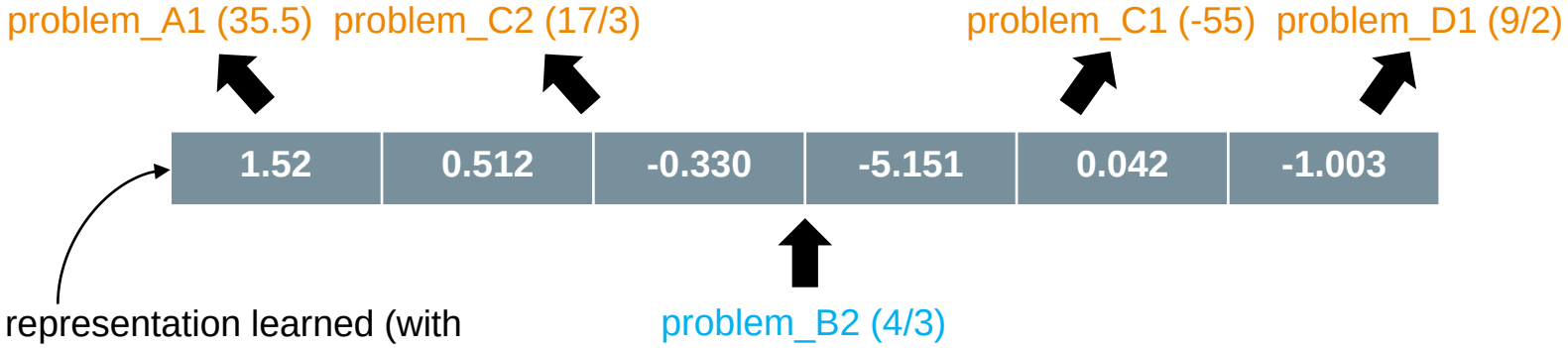


problem B2

-	Adding and Subtracting Fractions with Unlike Denominators	Multiplying Unit Fractions and Whole Numbers	Understanding Multiplying Fractions and Whole Numbers
# Student Answers	103,873	78,369	134,590
# Users	24,411	21,923	36,968

Our dataset consisting of three Khan Academy exercises

Student sequence of answers to problems in Fraction Subtraction exercise:

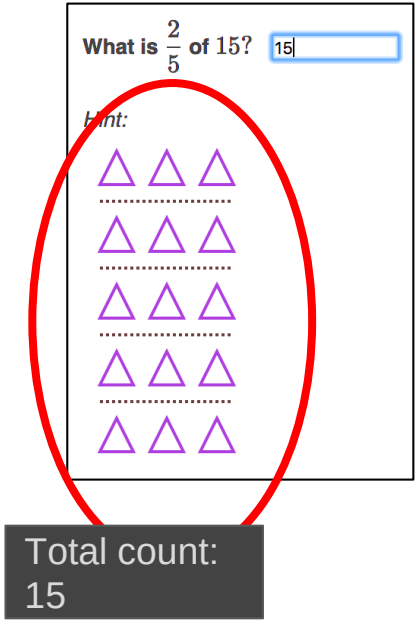


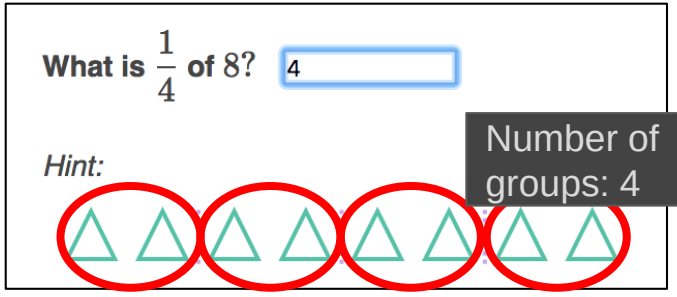
Vector representation learned (with backprop) which maximizes the accuracy of predicting the answers given across all context in which problem_B2 (4/3) is observed (across all student sequences)

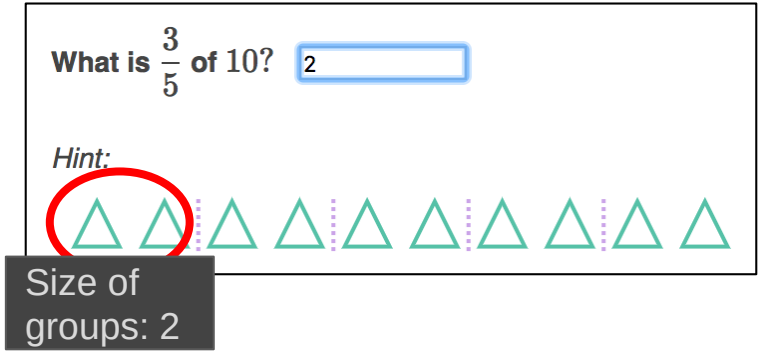
This algorithm is called a *Skip-gram* (word2vec) used in computational linguistics to model representations of words in large corpora (Mikolov et al., 2013)

Example problem answers from a cluster

30 clusters of 10 problem answers evaluated and hand tagged







- Is there meaningful similarity between these wrong answers?
- On average 46% of the 10 problem answers in the 30 clusters shared a misconception label

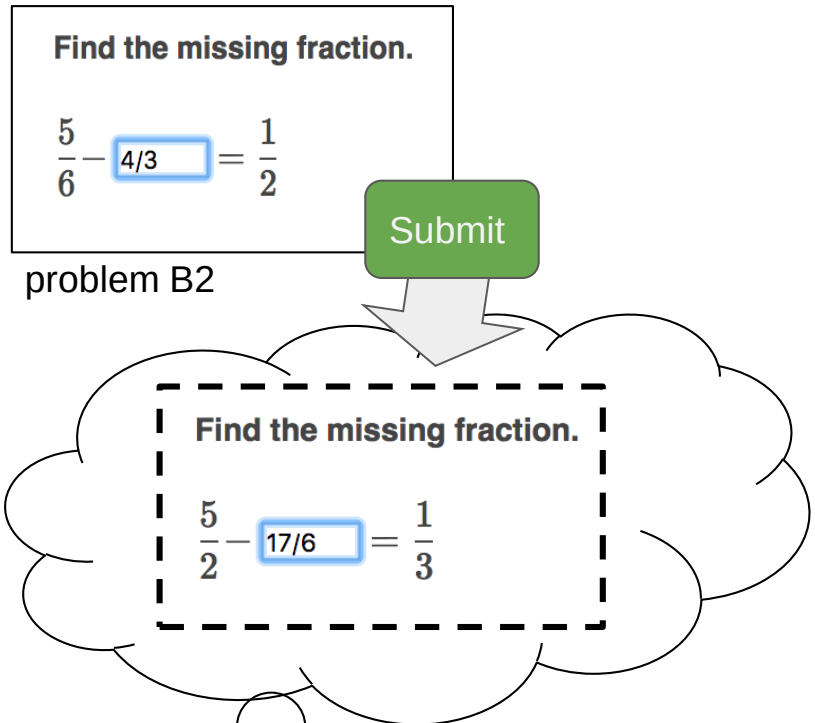
Three example answers from one cluster

(is the similarity related to a misconception?)

Designs adhering to guidance from theory

Smith et al. (1994) pressed for the importance of developing and refining students' conceptions through reflection and discussion (and **without direct confrontation**)

Possible intervention leverages similar wrong answers found **algorithmically**



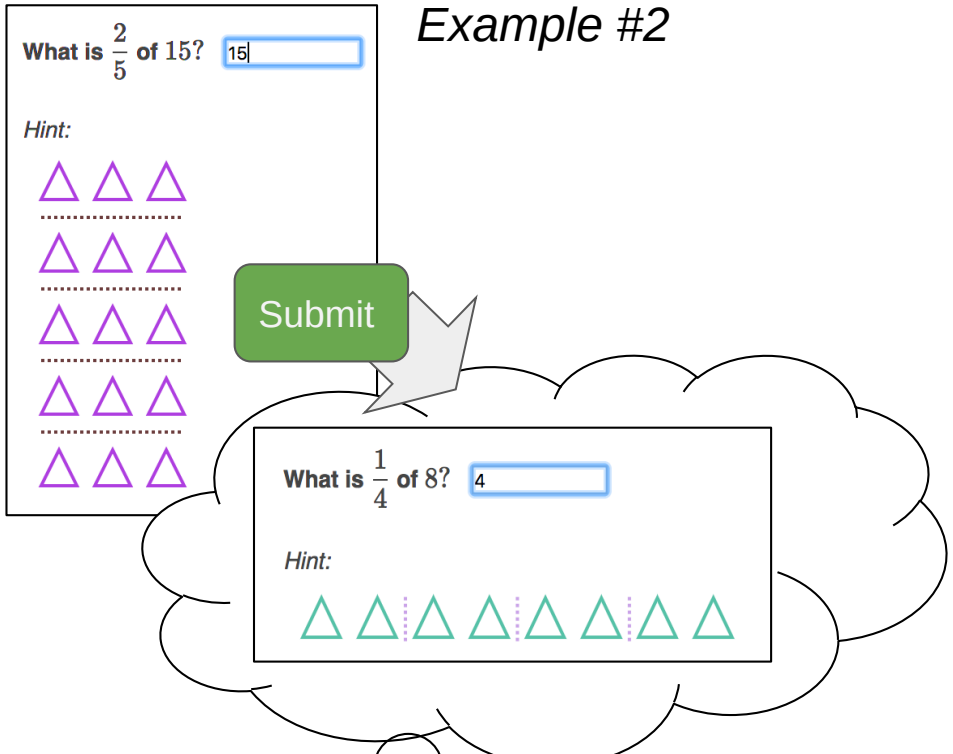
- You answered 4/3.
- Another student answered a question like this, **but made a mistake**. See if you can find the mistake and fix their answer for them.
- Do you want to change your answer?

Designs adhering to guidance from theory

Smith et al. (1994) pressed for the importance of developing and refining students' conceptions through reflection and discussion (and **without direct confrontation**)

Possible intervention leverages similar wrong answers found **algorithmically**

Example #2



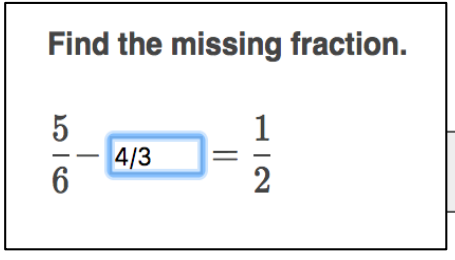
- You answered 15.
- Another student answered a question like this, **but made a mistake**. See if you can find the mistake and fix their answer for them.
- Do you want to change your answer?

Designs adhering to guidance from theory

Watkins et al. (2017) encourages student voicing of uncertainty as a mechanism to invite peer feedback, addressing incomplete understanding via co-construction

Distributed representation of misconceptions could aid in the orchestration of this online, matching students up with peers with complimentary constructions (e.g. wrong_answer1 + wrong_answer2 ≈ correct_answer in the vector space)

Student 1

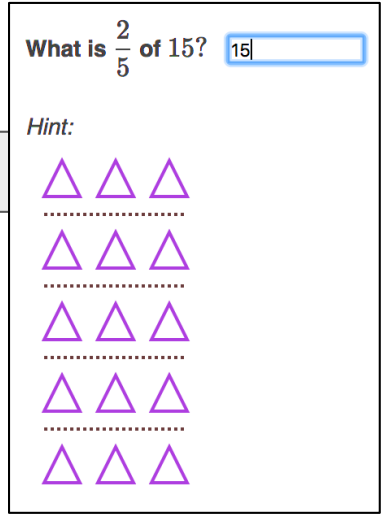


problem B2

Learners are paired together and engage in a discussion about their understanding

Dialogue

Student 2



Conclusions

- Distributed representations of wrong answers clustered together with 46% homogeneity with respect to hand tagged misconceptions
    - Confirms that the context of answers serves as a reasonably indicative signature of the misconception behind a wrong answer
- Answers did not cluster strictly by misconception, sometimes sharing a common unsuccessful problem solving strategy, instead
- The affordances of big data and representation learning algorithms can potentially scale and evaluate frameworks of learning from education theory
- Instructional design affordances of the model represent a new frontier for cooperation between the computational and learning sciences

Distributed Representation of Misconceptions

Thank You!

Questions?

Zachary A. Pardos
University of California at Berkeley

Scott Farrar
Khan Academy

John Kolb, Gao Xian Peh, Jong Ha Lee
University of California at Berkeley

Paper link: http://tiny.cc/**icls_distributed**

Acknowledgement: Thank you to the National Science Foundation (award #1547055), Alan Schoenfeld, and Khan Academy