

Fine-grained Cognitive Assessment based on Free-form Input for Math Story Problems

Bastiaan Heeren Johan Jeuring Sergey Sosnovsky Paul Drijvers Peter Boon
 Sietske Tacoma Jesse Koops Armin Weinberger Brigitte Grugeon-Allys
 Françoise Chenevotot-Quentin Jorn van Wijk Ferdinand van Walree



Proceedings of the European Conference on Technology Enhanced Learning (EC-TEL) 2018, pages 262–276, 2018

Motivation

Using ICT for **assessing** mathematics achievement:

- ▶ Targeted at 12-15 years old pupils
- ▶ Setting up algebraic expressions and equations, and simplifying and solving them

Detailed assessment of free-form answers to math story problems:

- ▶ Analysis of **intermediate steps**
- ▶ Determining the high-level **solution approach**
- ▶ Identification of **misconceptions**

Related assessment approaches

- ▶ Conventional assessment tests are usually unidimensional: instead, we focus on obtaining a detailed picture with strengths and weaknesses
- ▶ Easily verifiable input (e.g. multiple choice questions) may restrict the complexity of the tasks
- ▶ Structuring the interaction provides scaffolding to a pupil, which is less desirable in an assessment scenario
- ▶ Asking follow-up questions about intermediate steps may extend the assessment session

Our **research goal**: analysing free-form input to math story problems in an assessment setting

The screenshot shows the 'numworx' interface. The task is 'Magical trick?' with a story problem about a student's magic trick. The 'Your work' section shows a formula editor where the user has entered several algebraic expressions. A 'Submit' button is at the bottom.

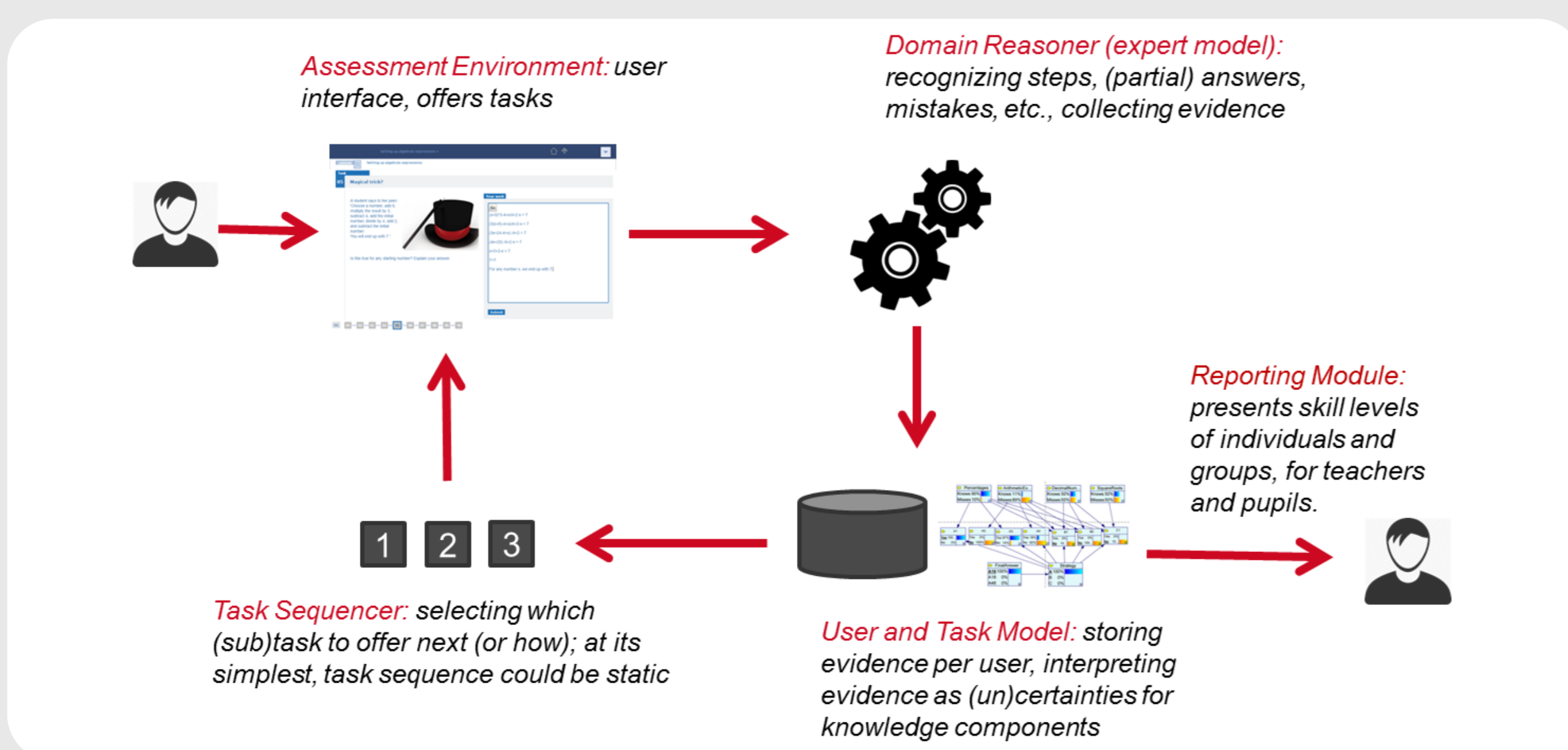
$((5+8)*3-4+5)/4+2-5 = 7$
 ▶ global, arithmetic

$5+8 = 13; 13*3 = 39;$
 $39-4 = 35; 35+5 = 40;$
 $40/4 = 10; 10+2 = 12;$
 $12-5 = 7$
 ▶ step-by-step, arithmetic

$(x+8)*3 = 3x+24$
 $3x+24-4 = 3x+20$
 $3x+20+x = 4x+20$
 $(4x+20)/4 = x+5$
 $x+5+2 = x+7$
 $x+7-x = 7$
 ▶ step-by-step, algebraic

$x+8*3-4+x/4+2-x$
 $= 2x+20/4+2-x$
 $= 2x+5+2-x$
 $= x+7$
 ▶ simplification mistakes

High-level architecture



Preliminary evaluation

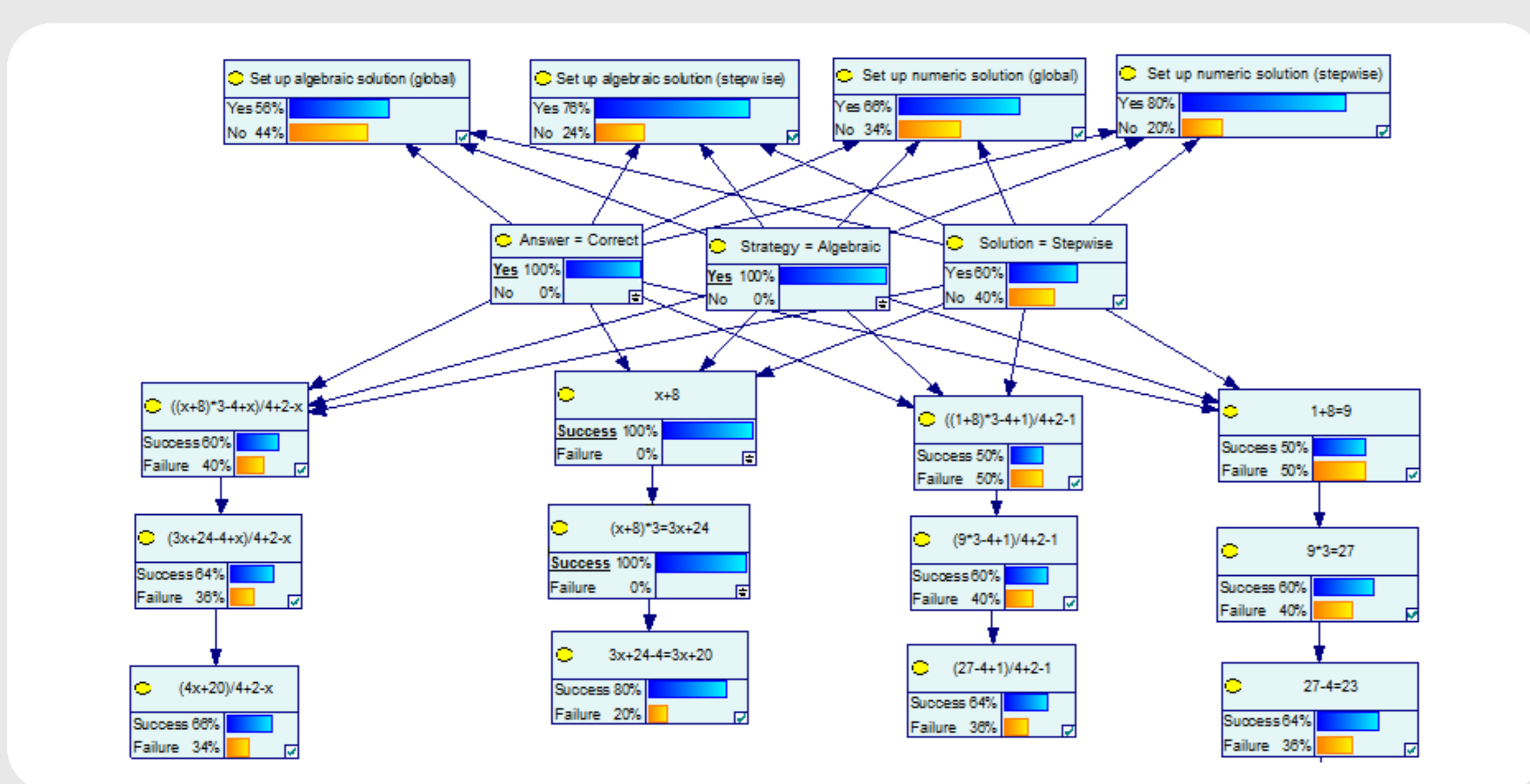
We tested the domain reasoner on a 'magical trick' dataset:

- ▶ 2956 free-form answers, collected between 2011 and 2015
- ▶ For nearly 80%, the solution approach was recognised

We organised two small-scale pilots:

- ▶ Pilots in Germany (N=19) and the Netherlands (N=22)
- ▶ Questionnaire, then 10 tasks, followed by a short survey
- ▶ Questionnaire focused on prior knowledge and usability
- ▶ Pupils think they did well in the test
- ▶ Again, the solution approach was recognised for nearly 80%

Example of a **task model**:



Conclusion

- ▶ We have developed a framework for **fine-grained cognitive assessment** of **free-form solutions** to **math story problems**
- ▶ The solution approach was recognised in nearly 80% of the answers for two small-scale pilot studies
- ▶ In the future, we will organise more extensive evaluations

The **Advise-Me project** (Automatic Diagnostics with Intermediate Steps in Mathematics Education) has received funding from the European Union's Erasmus+ Programme, Strategic Partnerships for school education for the development of innovation.

<http://advise-me.ou.nl>

✉ bastiaan.heeren@ou.nl

