Deep Inference for Automated Proof Tutoring?

Christoph Benzmüller[1,2], Dominik Dietrich[1], Marvin Schiller[1], Serge Autexier[1,3]

[1]Dept. of Computer Science, Saarland University, 66123 Saarbrücken, Germany
[2]Computer Laboratory, The University of Cambridge, Cambridge, CB3 0FD, UK
[3]DFKI, Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany

SFB 378: DIALOG

ΩMEGA for Proof Tutoring

- ΩMEGA [5] is a mathematical assistant system, developed at Saarland University.
- The DIALOG project [3] investigates natural language based tutoring of mathematical proofs supported by ΩMEGA.
- Student and DIALOG system interact to solve example problems; our running example is:

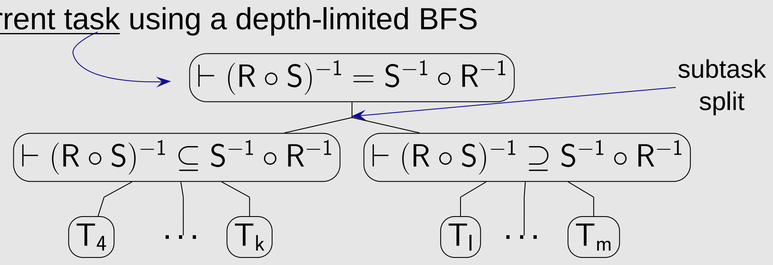
$$(R \circ S)^{-1} = S^{-1} \circ R^{-1}$$

- The task of the DIALOG system's proof asssesment module: proof step analysis (correctness, relevance and granularity) [2].
- Hypothesis investigated here: Does proof step analysis benefit from ΩMEGA's assertion application mechanism [7] which is based on the CORE calculus [1] (a deep inference variant).

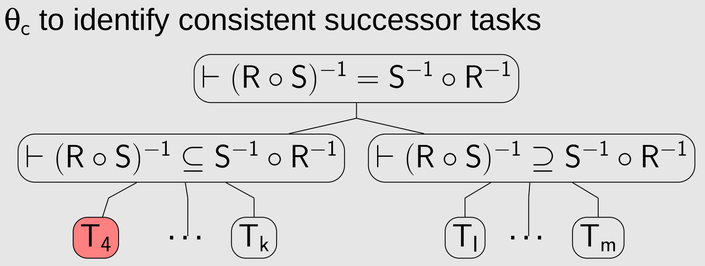
The Verification Process

Example student utterance: "Let $(x, y) \in (R \circ S)^{-1}$" – Task of the DIALOG system: Proof step analysis, in particular, correctness.

Expand the current task using a depth-limited BFS

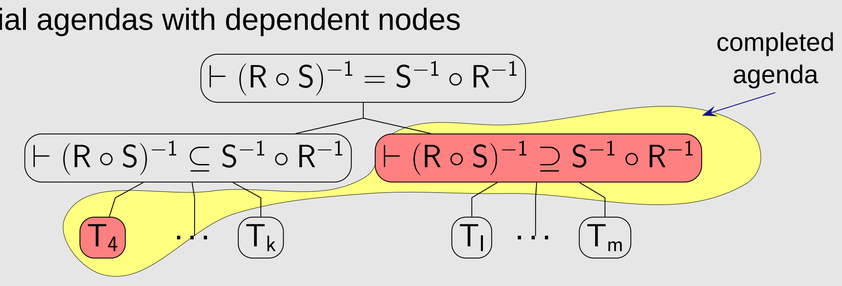


Apply the filter θ_c to identify consistent successor tasks

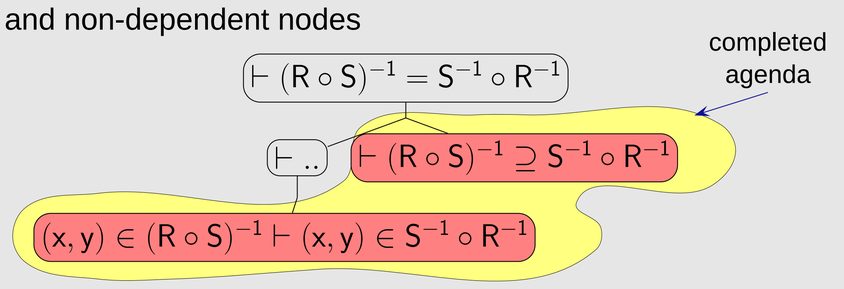


with T_4 : $(x, y) \in (R \circ S)^{-1} \vdash (x, y) \in S^{-1} \circ R^{-1}$

Complete partial agendas with dependent nodes



Prune rejected and non-dependent nodes



A consistent successor task has been found, so the step has been verified and we have a new cognitive proof state.

Example Dialog Fragment

S1: Let $(x, y) \in (R \circ S)^{-1}$.
T1: Good. correct
S2: It follows that $(y, x) \in (R \circ S)$.
T2: That's correct. correct
S3: Therefore $\exists z \in M : (y, z) \in R \wedge (z, x) \in S$.
T3: Yes, that's right. correct
S4: This implies: $\exists z \in M : (z, y) \in R^{-1} \wedge (x, z) \in S^{-1}$.
T4: Your deduction is correct. You are on a good way. correct
S5: It follows: $(x, y) \in S^{-1} \circ R^{-1}$.
T5: Yes. Therefore it is shown that every element from $(R \circ S)^{-1}$ already lies in $S^{-1} \circ R^{-1}$. correct

Corresponding Proof Reconstruction

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\text{S5: } (x,y) \in S^{-1} \circ R^{-1} \vdash \text{"—"}}\,\text{Close}}{\text{S4: } (z,y) \in R^{-1} \wedge (x,z) \in S^{-1} \vdash \text{"—"}}\,\text{Def.}\circ}{(y,z) \in R \wedge (x,z) \in S^{-1} \vdash \text{"—"}}\,\text{Def.}^{-1}}{\text{S3: } (y,z) \in R \wedge (z,x) \in S \vdash \text{"—"}}\,\text{Def.}^{-1}}{\text{S2: } (y,x) \in (R \circ S) \vdash \text{"—"}}\,\text{Def.}\circ}{\text{S1: } (x,y) \in (R \circ S)^{-1} \vdash (x,y) \in S^{-1} \circ R^{-1}}\,\text{Def.}^{-1}}{\vdash (R \circ S)^{-1} \subseteq S^{-1} \circ R^{-1}}\,\text{Def.}\subseteq \qquad \dfrac{(*)}{\texttt{t10:} \quad \vdash (R \circ S)^{-1} = S^{-1} \circ R^{-1}}\,\text{Def.}=$$

Sequent-style proof tree representation of ΩMEGA's proof automatically reconstructed for the given dialog fragment.
$(*)$ represents the proof obligation $\vdash (R \circ S)^{-1} \supseteq S^{-1} \circ R^{-1}$.

Observations

The student's proof fragment above can be verified with seven assertion level steps, which is still close to the student's original five steps. Compare these assertion level inferences to natural deduction inferences[a], for instance:

$$\text{Def}^{-1}\,\frac{(z,y) \in R^{-1} \wedge (x,z) \in S^{-1}}{(y,z) \in R \wedge (x,z) \in S^{-1}}$$

CORE

$$\wedge\text{I}\,\frac{\text{Def.}^{-1}\,\dfrac{\wedge\text{E}\,\dfrac{\varphi := (z,y) \in R^{-1} \wedge (x,z) \in S^{-1}}{(z,y) \in R^{-1}}}{(y,z) \in R} \qquad \wedge\text{E}\,\dfrac{\varphi}{(x,z) \in S^{-1}}}{(y,z) \in R \wedge (x,z) \in S^{-1}}$$

Natural Deduction

[a]Natural Deduction was the basis of an earlier version of the ΩMEGA system

Evaluation

- Application of proof assessment module to 17 tutorial dialogs from a Wizard-of-Oz experiment [3]
- 144 proof steps over 17 sessions
- Proof search limited to depth 4

correctly rejected:	28	19 %
correctly verified :	113	79 %
wrongly verified :	0	0 %
wrongly rejected :	3	2 %

Conclusion

- Proof step analysis, in particular, correctness analysis, benefits from ΩMEGA's assertion level proofs (which are close to human proofs).
- Using simple BFS proof search at assertion level we were able to correctly classify 95.9% of the steps in the 17 tutorial dialogs.

References

[1] S. Autexier. The CORE calculus. Proc. CADE-20, LNAI 3632, 2005.

[2] C. Benzmüller and Q.B. Vo. Mathematical domain reasoning tasks in natural language tutorial dialog on proofs. Proc. AAAI-05. 2005.

[3] C. Benzmüller et al. Natural language dialog with a tutor system for mathematical proofs. In *Cognitive Systems*. LNAI 4429, 2007.

[4] D. Dietrich and M. Buckley. Verification of Proof Steps for Tutoring Mathematical Proofs. Proc. AIED 2007.

[5] J. Siekmann et al. Computer supported mathematics with ΩMEGA. *J. Applied Logic*, 4(4):533–559, 2006.

[6] S. Autexier and D. Dietrich. Synthesizing Proof Planning Methods and Ω-ANTS Agents from Mathematical Knowledge. Proceedings of MKM'06. LNAI 4108, 2006.

[7] D. Dietrich. The tasklayer of the ΩMEGA system. Diploma thesis, Universität des Saarlandes, Saarbrücken, Germany, 2006.