Granularity Judgments in Proof Tutoring

SFB 378 Ressourcenadaptive Kognitive Prozesse

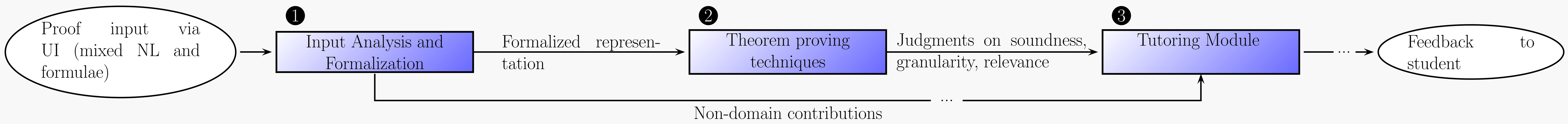
Marvin Schiller and Christoph Benzmüller

Saarland University, D-66041 Saarbrücken, Germany

Universität des Saarlandes

SFB 378 project Dialog

The SFB 378 project Dialog [2] investigates natural tutorial dialog between a student and an assistance system for mathematics.



Granularity in Mathematical Dialogs

Granularity: size of a proof step w.r.t its argumentative complexity.
Excerpts from proofs of $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$:

Let x be an element of $A \cap (B \cup C)$, then $x \in A$ and $x \in B \cup C$. This means that $x \in A$, and either $x \in B$ or $x \in C$.

small steps, more detail

Let x be an element of $A \cap (B \cup C)$. This means that $x \in A$, and either $x \in B$ or $x \in C$.

bigger steps, less detail

↪ Same line of reasoning, but different granularities

Does granularity play a role in tutorial dialogs on proofs?

- We collected a corpus of dialogs in Wizard-of-Oz experiments
- The wizards annotated all student proof steps w.r.t. granularity w.r.t. three categories, *too detailled*, *appropriate*, *too coarse-grained*.

Observation

On average, the wizards identified 1.92 utterances as non-appropriate w.r.t granularity (out of an average of 25 dialog contributions per student).

Wizard-of-Oz subject (left picture) and the wizard (right picture, on the right) during the experiment.

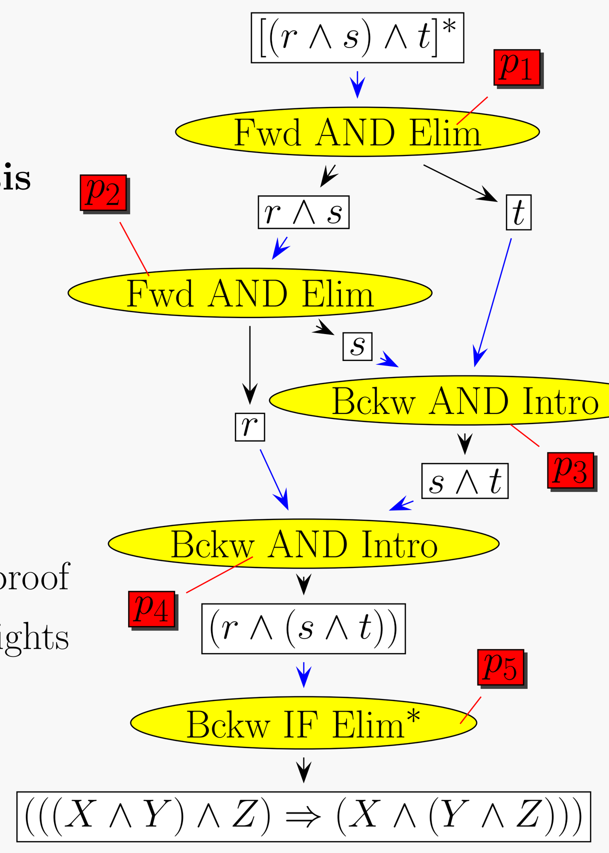
Framework and Calculi for Granularity Analysis

Generic Framework for Granularity Analysis
Parameterizable over

- proof calculus
- weighting
- repetition effect

Approach

1. Student proof step is expanded into calculus level proof
2. Nodes of proof tree/proof graph are assigned weights $p_1, p_2, \ldots, p_n$
3. Size of formalized proof step = weighted sum



What are good proof calculi for granularity analysis?
...preferably human-oriented, cognitively adequate calculi.

Hypothesis: Size of formalized proof step in a human-oriented calculus can serve as an indicator for granularity.

Two candidate calculi:

Gentzen's natural deduction (ND) calculus [3]

- "[...] I wished to construct a formalism that comes as close as possible to actual reasoning." [3]

$$\dfrac{Man(A) \quad \dfrac{\forall x(Man(x) \Rightarrow Mortal(x))}{Man(A) \Rightarrow Mortal(A)}\forall Elim}{Mortal(A)} \Rightarrow Elim$$

"Psychology of Proof" (PSYCOP) calculus [5]

- quantifier-free formula representations
- decision procedure for proof search (incomplete, but cognitive adequacy supported by emprical studies)

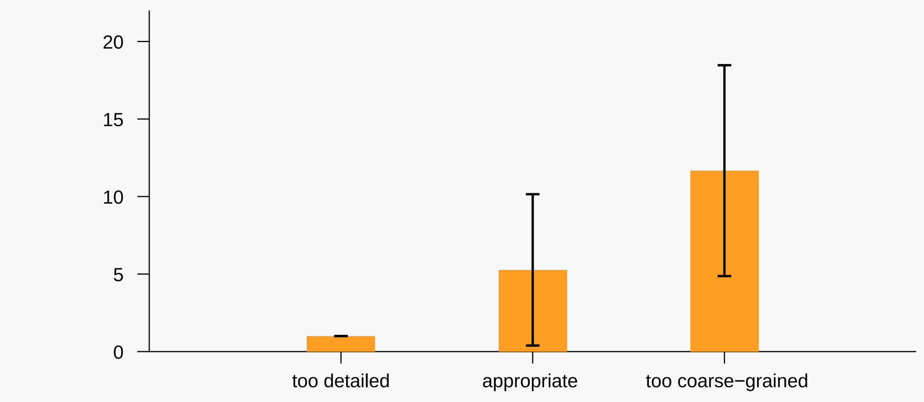
$$\dfrac{Man(A) \quad Man(x) \Rightarrow Mortal(x)}{Mortal(A)} Backward \Rightarrow Elim$$

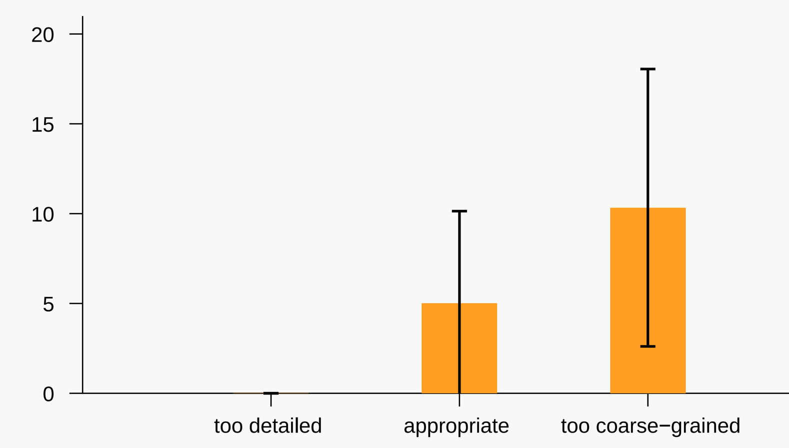
Evaluation

- We instantiated our framework with the ND and PSYCOP calculi (with equal weights for all inference rules),
- applied these framework instances to student proof steps from the corpus,
- and related the resulting proof size figures to the tutors' granularity ratings.

Results

1. Average calculus level proof sizes indeed reflect the granularity level identified by the tutor.
2. However, standard deviations indicate that sizes of calculus level proofs vary greatly within each group; no classification possible based on calculus level proof length alone.





Average number of calculus level proof steps in ND (left) and PSYCOP (right) that constitute a student's proof step for twenty steps, grouped by their granularity level as identified by the tutors. Bars indicate corresponding standard deviations.

Discussion

- Students often used rewriting/assertion level [4]/deep inference proof steps in the experiment; these are not modelled appropriately in ND or PSYCOP.
- Therefore, ongoing work: building an instance of granularity analysis framework based on Ωmega-CoRe [1] calculus,
- Determine and investigate calculus-specific weights for granularity analysis.
- Planning of another experiment for specifically investigating granularity phenomena.
- Granulariy annotations by wizards reflect subjectivity - this suggests coupling the presented framework with student and teaching models.

References

[1] S. Autexier. *Hierarchical Contextual Reasoning.* PhD thesis, Saarland University, Saarbrücken, Germany, 2003.

[2] C. Benzmüller et al. Tutorial dialogs on mathematical proofs. In *Proc. of the IJCAI Workshop on Knowledge Representation and Automated Reasoning for E-Learning Systems*, pages 12–22, Acapulco, 2003.

[3] G. Gentzen. Untersuchungen über das logische Schliessen. *Mathematische Zeitschrift*, 39:176–210, 405–431, 1934.

[4] X. Huang. *Human Oriented Proof Presentation: A Reconstructive Approach.* Phd thesis, Universität des Saarlandes, Saarbrücken, Germany, 1994.

[5] L. J. Rips. *The Psychology of Proof.* MIT Press, Cambridge, MA, 1994.