

Granularity-Adaptive Proof Presentation^a

Marvin Schiller[†] and Christoph Benzmüller^{*}



[†]German Research Center for Artificial Intelligence (DFKI), Bremen, Germany

^{*}International University in Germany, Bruchsal, Germany

marvin.schiller@dfki.de, c.benzmueller@googlemail.com



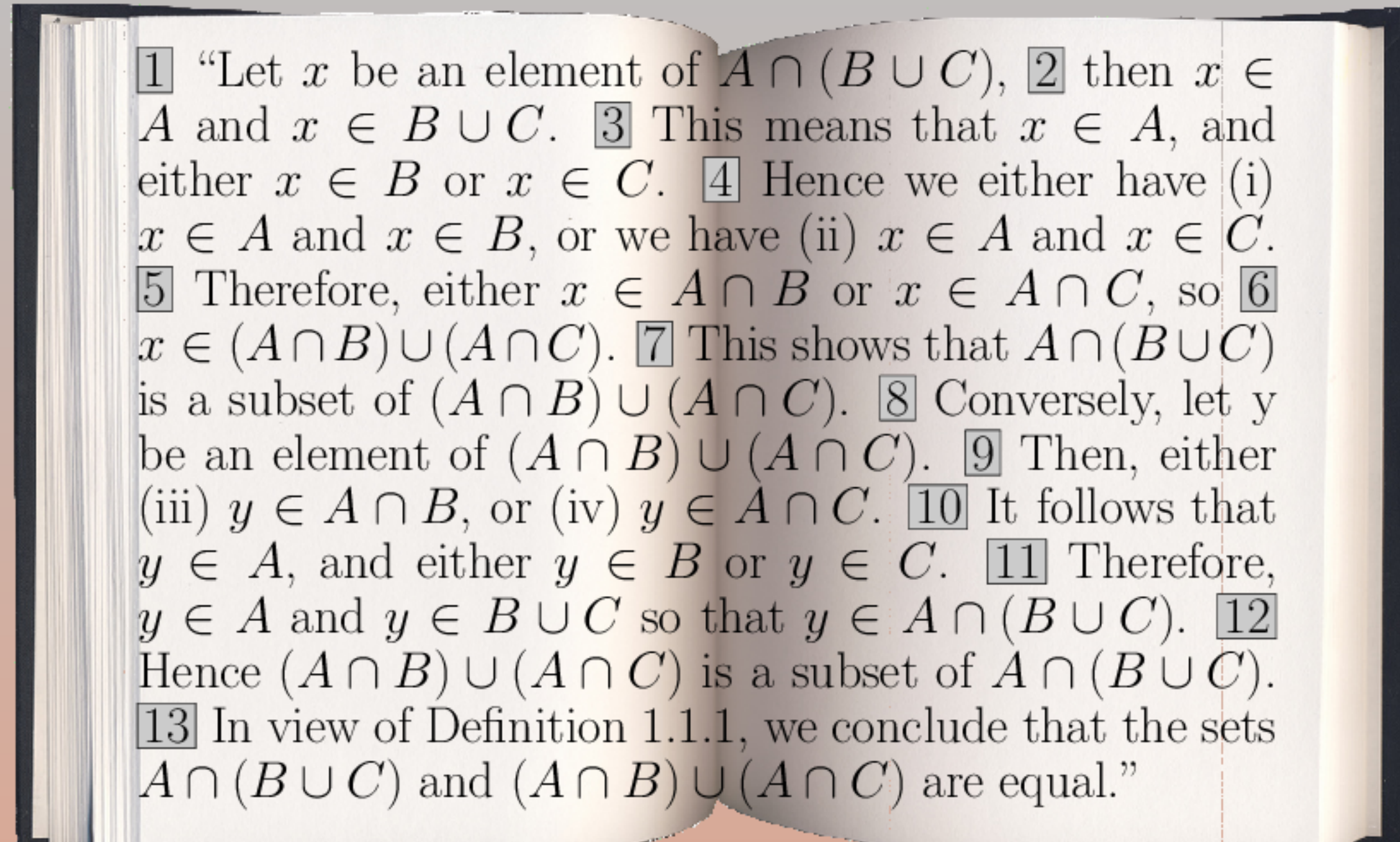
Studienstiftung
des deutschen Volkes



UNIVERSITÄT
DES
SAARLANDES

^aThis work was supported by the SFB 378 Project DIALOG and the German National Merit Foundation.

Granularity



Proof of $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, reproduced from [2]

Granularity matters in mathematics. For example, in introductory textbooks intermediate proof steps are often skipped, when this seems appropriate, e.g.

9 [...] either (iii) $y \in A \cap B$, or (iv) $y \in A \cap C$.
 $y \in A \wedge y \in B$ or $y \in A \cap C$
 $y \in A \wedge y \in B$ or $y \in A \wedge y \in C$
10 It follows that $y \in A$, and either $y \in B$ or $y \in C$.

$\left. \begin{array}{l} y \in A \wedge y \in B \text{ or } y \in A \cap C \\ y \in A \wedge y \in B \text{ or } y \in A \wedge y \in C \end{array} \right\} \begin{array}{l} 2 \times \text{Def.}\cap, \\ \text{Distributivity} \end{array}$

How can the notion of appropriate step size be captured and implemented within an automated environment?

Objectives

1. Model/represent different granularities
2. Adapt proof in formal systems (here, Ω MEGA [1]) for output at suitable granularity level
3. Learn granularity from empirical samples

Approach

1. Chunk assertion level proof steps according to granularity classifier (class labels: too-big, too-small, appropriate)
2. Granularity classifier as 'good old' AI expert system, trained via machine learning from samples (e.g. ruleset classifiers learnt via C5.0 [4, 3])

Granularity Criteria

Mastered vs. unmastered concepts with respect to a student model (continually updated)

What concepts? - a feature value for each concept

Introduced hypotheses/subgoals

What theory do the employed concepts belong to?

Are the employed concepts mentioned verbally?

Assertion level proofs (where each inference application is justified by a mathematical fact, such as a definition, lemma or theorem) are well-suited to obtain such granularity-relevant information.

Assertion Level Proof

	$x \in S \vdash x \in S$	$y \in T \vdash y \in T$
DEFU (8)	$\frac{}{(x \in (A \cap B) \vee x \in (A \cap C)) \vdash x \in S}$	$\frac{}{(y \in A \wedge y \in (B \cup C)) \vdash y \in T}$ DEFU (15)
DEFU (7)	$\frac{}{(x \in (A \cap B) \vee x \in A \wedge x \in C) \vdash x \in S}$	$\frac{}{(y \in A \wedge (y \in B \vee y \in C)) \vdash y \in T}$ DEFU (14)
DEFU (6)	$\frac{}{(x \in A \wedge x \in B \vee x \in A \wedge x \in C) \vdash x \in S}$	$\frac{}{(y \in A \wedge y \in B \vee y \in A \wedge y \in C) \vdash y \in T}$ DISTR (13)
DISTR (5)	$\frac{}{(x \in A \wedge (x \in B \vee x \in C)) \vdash x \in S}$	$\frac{}{(y \in A \wedge y \in B \vee y \in (A \cap C)) \vdash y \in T}$ DEFU (12)
DEFU (4)	$\frac{}{(x \in A \wedge x \in (B \cup C)) \vdash x \in S}$	$\frac{}{(y \in (A \cap B) \vee y \in (A \cap C)) \vdash y \in T}$ DEFU (11)
DEFU (3)	$\frac{}{(x \in (A \cap (B \cup C))) \vdash x \in S}$	$\frac{}{(y \in S) \vdash y \in T}$ DEFU (10)
DEFU (2)	$\frac{}{\vdash (A \cap (B \cup C)) \subseteq S}$	$\frac{}{\vdash ((A \cap B) \cup (A \cap C)) \subseteq T}$ DEFU (9)
DEF EQ (1)	$\frac{}{\vdash \underbrace{(A \cap (B \cup C))}_T = \underbrace{((A \cap B) \cup (A \cap C))}_S}$	

Classifier 1

Adaptive Proof Presentation

Classifier 1

Classifier 2

13 We show that $((A \cap B) \cup (A \cap C) \subseteq A \cap B \cup C)$ and $(A \cap B \cup C \subseteq (A \cap B) \cup (A \cap C))$...because of definition of equality
1 We assume $x \in A \cap B \cup C$ and show $x \in (A \cap B) \cup (A \cap C)$
2 Therefore, $x \in A \wedge x \in B \cup C$
3 Therefore, $x \in A \wedge (x \in B \vee x \in C)$
4 Therefore, $x \in A \wedge x \in B \vee x \in A \wedge x \in C$
5 Therefore, $x \in A \cap B \vee x \in A \wedge x \in C$
6 Therefore, $x \in A \cap B \vee x \in A \cap C$
7 We are done with the current part of the proof (i.e.; to show that $x \in (A \cap B) \cup (A \cap C)$). [It remains to be shown that $(A \cap B) \cup (A \cap C) \subseteq A \cap B \cup C$]
: 1 [alternatively 2, etc.]
12 This finishes the proof. Q.E.D.

Proof presentation generated from assertion level proof, which matches the example textbook proof in step size (modulo some re-ordering of steps).

1 We assume $y \in (A \cap B) \cup (A \cap C)$ and show $y \in A \cap B \cup C$
8 We assume $y \in (A \cap B) \cup (A \cap C)$ and show $y \in A \cap B \cup C$
9 Therefore, $y \in A \cap B \vee y \in A \cap C$
10 Therefore, $y \in A \wedge (y \in B \vee y \in C)$
11 Therefore, $y \in A \wedge y \in B \cup C$

The presentation algorithm classifies to-be-presented steps and skips intermediate steps if they are too small as such. Using different classifiers allows to vary the granularity. Granularity classifiers are sensitive to changes in user knowledge via input from the user model.

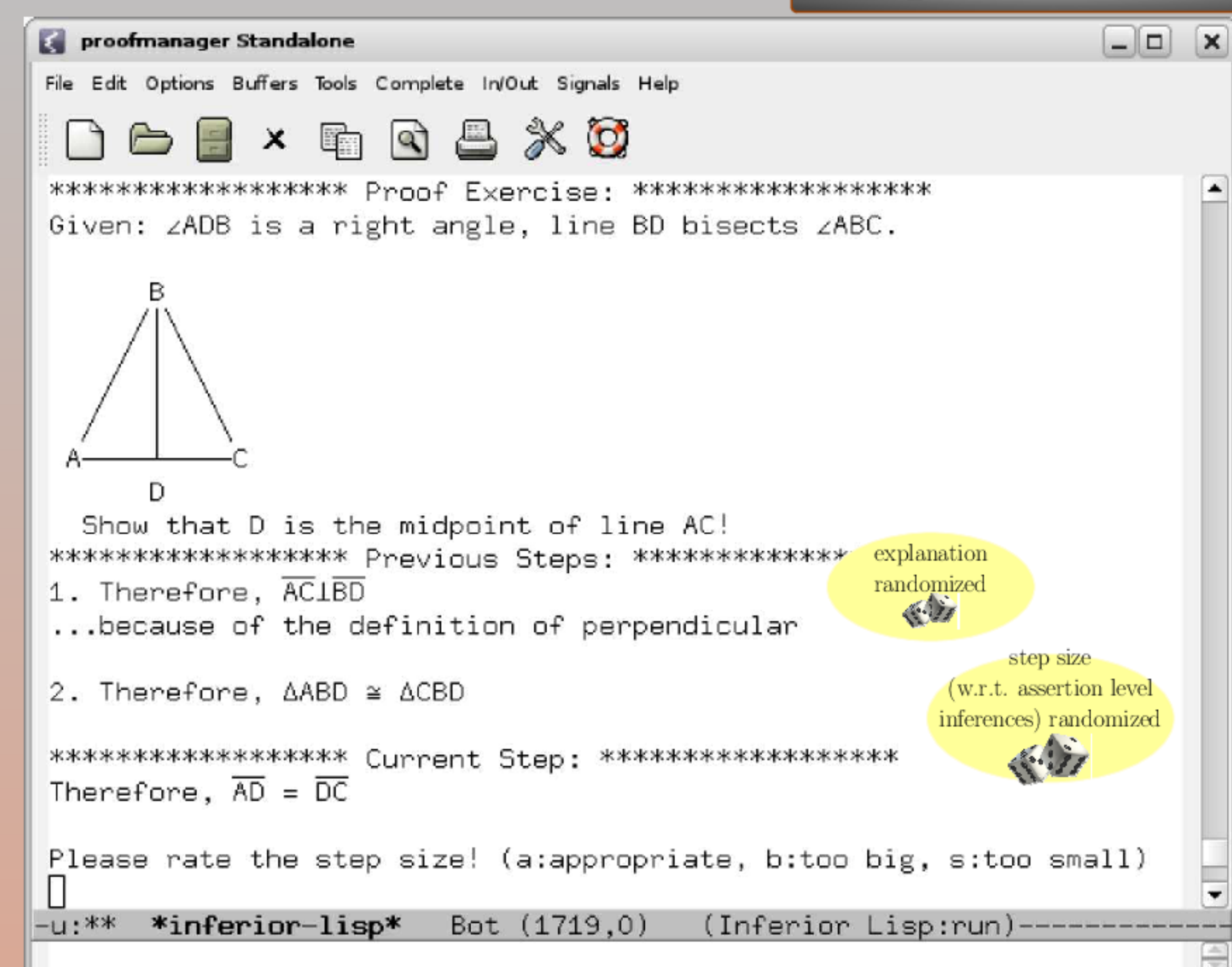
Classifier 1

1) conceptsunique $\in \{0, 1\} \wedge \text{def-EQ}=0 \wedge \text{verb}=\text{true} \Rightarrow \text{step-too-small}$
2) hypintro=0 $\wedge \text{def-EQ}=0 \wedge \text{def-U}=0 \wedge \text{verb}=\text{true} \Rightarrow \text{step-too-small}$
3) conceptsunique $\in \{2, 3, 4\} \wedge \text{def-U} \in \{1, 2, 3\} \Rightarrow \text{step-too-big}$
4) hypintro $\in \{1, 2, 3, 4\} \wedge \text{conceptsunique} \in \{2, 3, 4\} \Rightarrow \text{step-too-big}$
5) unmasteredconceptsunique=0 $\wedge \text{total} \in \{0, 1, 2\} \wedge \text{def-U} \in \{1, 2\} \wedge \text{close}=\text{false} \Rightarrow \text{step-too-small}$
6) def-EQ $\in \{1, 2\} \wedge \text{verb}=\text{false} \Rightarrow \text{step-too-big}$
7) def-EQ $\in \{1, 2\} \wedge \text{verb}=\text{true} \Rightarrow \text{step-appropriate}$
8) def-EQ=0 $\wedge \text{verb}=\text{false} \Rightarrow \text{step-appropriate}$
9) $_ \Rightarrow \text{step-appropriate}$

Classifier 2

$_ \Rightarrow \text{step-appropriate}$

Empirical Evaluations



The study environment allows to sample granularity judgments by experts and evaluate the performance of the learnt classifiers.

Recent Evaluation Results

- 2 experiments with 2 expert tutors (using different exercises in naive set theory, relations, topology)

	Tutor 1	Tutor 2
Steps annotated:	135	207
Perf. learnt classifier ¹	86.7%	68.9%
-mean correct	$\kappa=0.68$	$\kappa=0.47$
- κ		
Interrater reliability ²	$\kappa=0.37$	

¹best rule-based classifier, evaluated on full dataset using 10-fold cross validation
²on common subset of 108 steps

Discussion

- We have explored a classification-based approach to granularity.
- Assertion level proofs provide useful information for the classification task, and for natural language output.
- Further work consists in empirical investigations, examining the usefulness of different granularity criteria and individual differences between human experts.

References

- [1] S. Autexier, C. Benzmüller, D. Dietrich, A. Meier, and C.-P. Wirth. A generic modular data structure for proof attempts alternating on ideas and granularity. In M. Kohlhase, editor, Proc.of MKM'05, volume 3863 of LNCS. Springer, 2006.
- [2] R. G. Bartle and D. Sherbert. Introduction to Real Analysis. Wiley, 2 edition, 1982.
- [3] J. R. Quinlan. C4.5: Programs for Machine Learning. Morgan Kaufmann, 1993.
- [4] RuleQuest Research. Data mining tools see5 and c5.0. <http://www.rulequest.com/see5-info.html>, retrieved April 15, 2009.