Granularity in Mathematical Dialog

Marvin Schiller
Saarland University

schiller@ags.uni-sb.de.

Theorema-Ultra-Omega'05 Workshop
November 15th, 2005
Saarland University

Overview

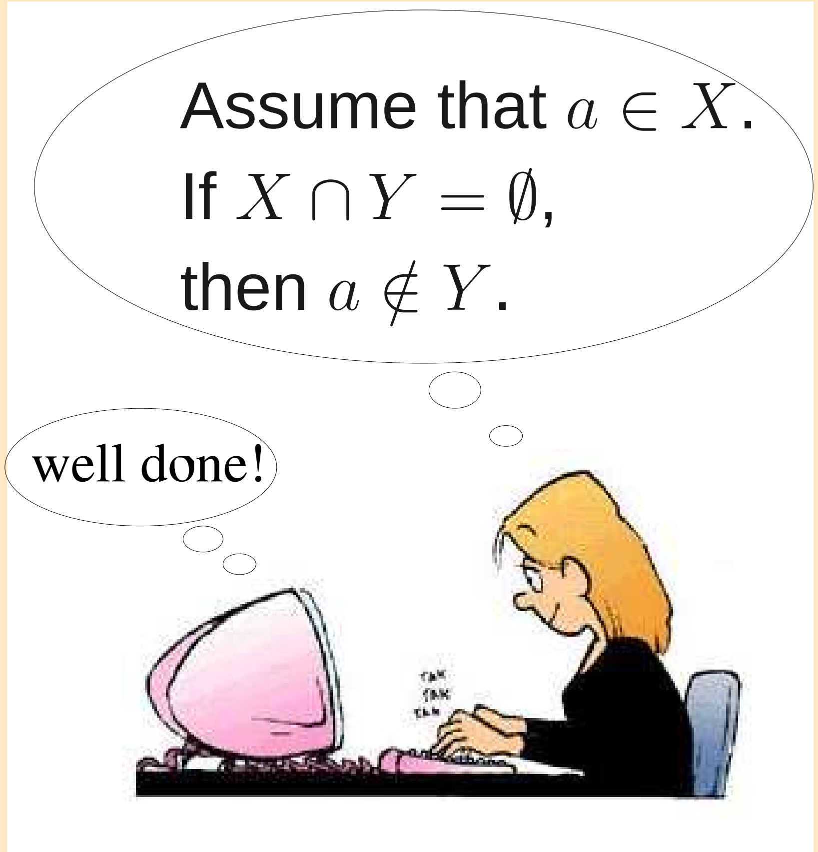
- Background: Granularity in mathematics and mathematical dialogs
- Approach and mechanization
- Analysis & Evaluation
- Observations
- Further Directions

Objective

Intelligent Tutoring/Dialog Systems for Mathematics: help with solving maths problems/learning maths.

Tasks for math. dialog systems:

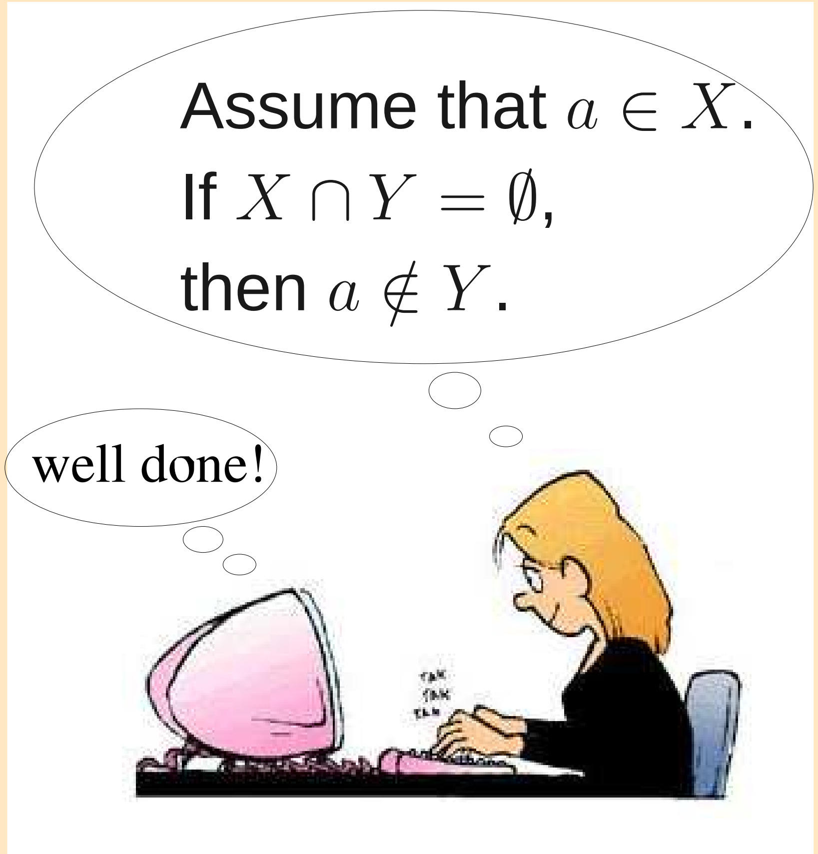
- Natural language analysis
- Mathematical domain reasoning
- Dialog management
- Output generation and verbalization



Tasks for Math. Domain Reasoning

Tasks (Benzmüller & Vo 2005):

- Analyze mathematical accuracy/correctness
- Analyze granularity
- Analyze relevance



Granularity ...in the corpus

Granularity ... the size/argumentative complexity of a proof step

Granularity factors:

- abstraction
- explicitness/underspecification
- cognitive effort

Granularity in ’05 Wizard-of-Oz Study

- Task for the participants: Collaboratively solve mathematical proofs with a mathematical tutoring system (domain: relations).
- Task for the wizards: Categorization of mathematical dialog contributions:

student]	Let $(x, y) \in (R \circ S)^{-1}$
tutor]	Correct. Good start!

correct	appropriate	relevant

Aspects I

- abstraction

student]	let $(x, y) \in (R \cup S) \circ T$, then $(x, z) \in (R \cup S) \wedge (z, y) \in T$
	(...) then $(x, z) \in R \vee (x, z) \in S$
tutor]	This statement is true.
student]	what can be concluded from $(A \vee B) \wedge C$?
tutor]	Then for example it holds $(A \wedge C) \vee (B \wedge C)$
student]	then holds $((x, z) \in R \wedge (z, y) \in T) \vee ((x, z) \in S \wedge (z, y) \in T)$

- explicitness/underspecification
- cognitive effort

Aspects (II)

- abstraction
- explicitness/underspecification

student 13]	$(R \cup S) \circ T$ thus is
...	
student 19]	$(R \circ T) \cup (S \circ T)$ thus is (...)
tutor]	Correct. Can you also indicate according to which law you have transfomed input 13 to the current input 19?
student 20]	“distributivity law”

- cognitive effort

Aspects(III)

- abstraction
- explicitness/underspecification
- cognitive effort

student]	$(x, y) \in (R \circ S)^{-1}$
tutor]	Now try to draw inferences from that!
student]	$(x, y) \in S^{-1} \circ R^{-1}$
tutor]	One cannot directly deduce that.
	You need some intermediate steps!

Approach

Mechanize granularity ratings with thm. proving techniques:

- Use ΩMEGA framework.
- Hypothesis: granularity level of a mathematical statement is related to number of inference steps required for its justification.
- Calculi: Gentzen's ND (Gentzen 1934) and "Psychology of Proof" (Rips 1994).
- Granularity analysis framework for proofs.
- Evaluation: compare mechanical classification to expert's ratings.

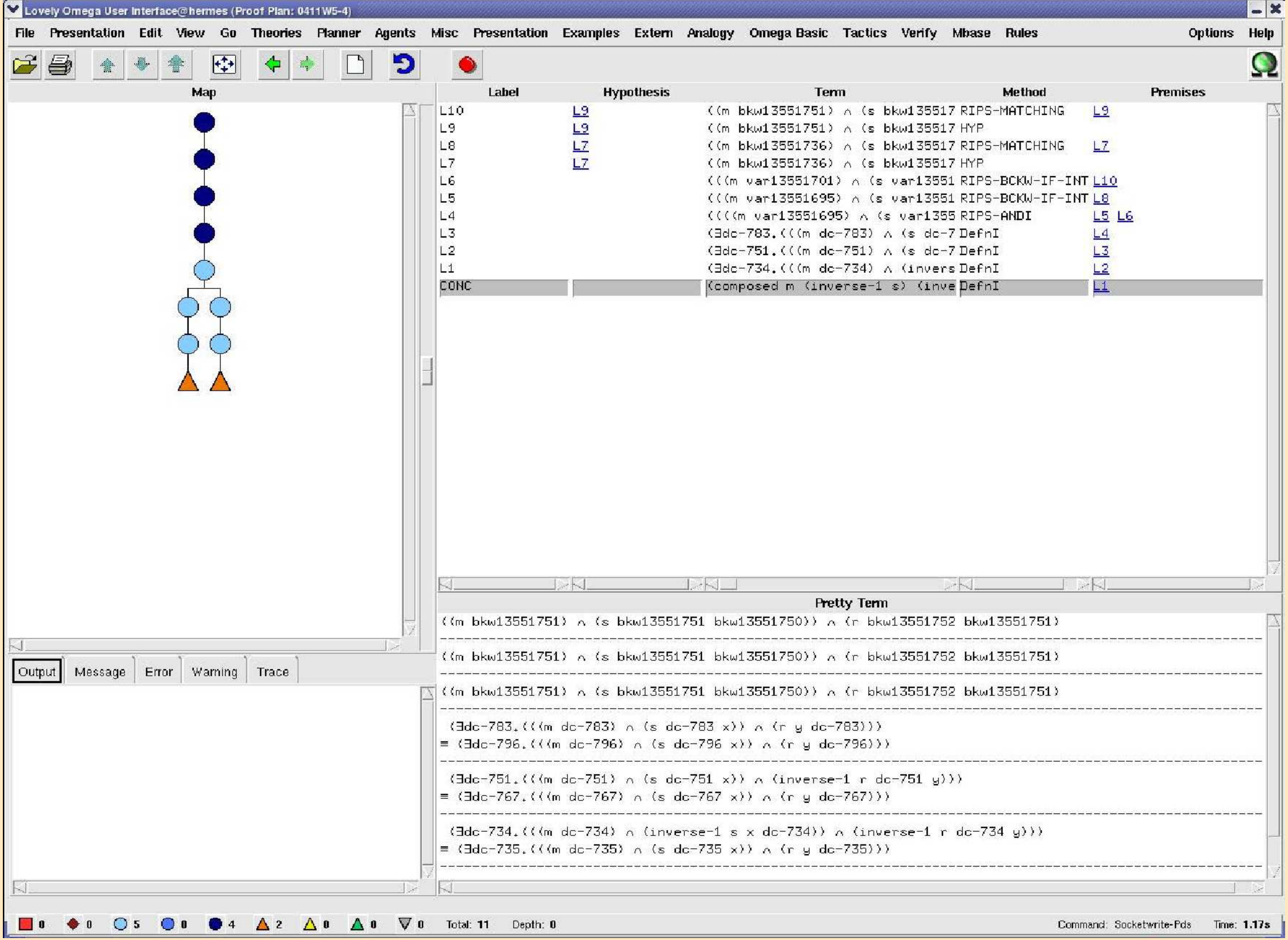
Granularity Evaluation

Example:

student]	$(x, y) \in (S^{-1} \circ R^{-1}) \Leftrightarrow \exists z[(z, x) \in S \wedge (y, z) \in R]$			
tutor]	This is correct!	correct	too coarse-grained	relevant

... with the help of a parser:

```
(equiv ((composed M (inverse-1 S)
                    (inverse-1 R)) x y)
       (exists (lam (z a)
            (and (and (M z) (S z x)) (R y z)))))
```


Lovely Omega User Interface@hermes (Proof Plan: 0411W5-4)
File Presentation Edit View Go Theories Planner Agents Misc Presentation Examples Extern Analogy Omega Basic Tactics Verify Mbase Rules Options Help
Map
Label Hypothesis Term Method Premises
L10 L9 ((m bkw13551751) ∧ (s bkw135517 RIPS-MATCHING L9
L9 L9 ((m bkw13551751) ∧ (s bkw135517 HYP
L8 L7 ((m bkw13551736) ∧ (s bkw135517 RIPS-MATCHING L7
L7 L7 ((m bkw13551736) ∧ (s bkw135517 HYP
L6 (((m var13551701) ∧ (s var13551 RIPS-BCKW-IF-INT L10
L5 (((m var13551695) ∧ (s var13551 RIPS-BCKW-IF-INT L8
L4 ((((m var13551695) ∧ (s var1355 RIPS-ANDI L5 L6
L3 (∃dc-783.(((m dc-783) ∧ (s dc-7 DefnI L4
L2 (∃dc-751.(((m dc-751) ∧ (s dc-7 DefnI L3
L1 (∃dc-734.(((m dc-734) ∧ (invers DefnI L2
CONC (composed m (inverse-1 s) (inve DefnI L1
Pretty Term
((m bkw13551751) ∧ (s bkw13551751 bkw13551750)) ∧ (r bkw13551752 bkw13551751)
((m bkw13551751) ∧ (s bkw13551751 bkw13551750)) ∧ (r bkw13551752 bkw13551751)
((m bkw13551751) ∧ (s bkw13551751 bkw13551750)) ∧ (r bkw13551752 bkw13551751)
(∃dc-783.(((m dc-783) ∧ (s dc-783 x)) ∧ (r y dc-783)))
≡ (∃dc-796.(((m dc-796) ∧ (s dc-796 x)) ∧ (r y dc-796)))
(∃dc-751.(((m dc-751) ∧ (s dc-751 x)) ∧ (inverse-1 r dc-751 y)))
≡ (∃dc-767.(((m dc-767) ∧ (s dc-767 x)) ∧ (r y dc-767)))
(∃dc-734.(((m dc-734) ∧ (inverse-1 s x dc-734)) ∧ (inverse-1 r dc-734 y)))
≡ (∃dc-735.(((m dc-735) ∧ (s dc-735 x)) ∧ (r y dc-735)))
Output Message Error Warning Trace
0 0 5 0 4 2 0 0 0 Total: 11 Depth: 0
Command: Socketwrite-Pds Time: 1.17s

Sample Analysis

A: $(x, y) \in (S^{-1} \circ R^{-1}) \Leftrightarrow \exists z[(z, x) \in S \wedge (y, z) \in R]$

B: $\forall x \forall y[\exists z[(y, z) \in R \wedge (z, x) \in S] \rightarrow (y, x) \in (R \circ S)]$

C: therefore it follows: $(x, y) \in (S^{-1} \circ R^{-1}) \rightarrow (y, x) \in (R \circ S)$

	Statement A	Statement B	Statement C
PSYCOP			
[Gentzen34]			
Tutor			

Number of justifying proof steps for PSYCOP and Gentzen's NK.

Sample Analysis

A: $(x, y) \in (S^{-1} \circ R^{-1}) \Leftrightarrow \exists z[(z, x) \in S \wedge (y, z) \in R]$

B: $\forall x \forall y[\exists z[(y, z) \in R \wedge (z, x) \in S] \rightarrow (y, x) \in (R \circ S)]$

C: therefore it follows: $(x, y) \in (S^{-1} \circ R^{-1}) \rightarrow (y, x) \in (R \circ S)$

	Statement A	Statement B	Statement C
PSYCOP	5	2	10
[Gentzen34]			
Tutor			

Number of justifying proof steps for PSYCOP and Gentzen's NK.

Sample Analysis

A: $(x, y) \in (S^{-1} \circ R^{-1}) \Leftrightarrow \exists z[(z, x) \in S \wedge (y, z) \in R]$

B: $\forall x \forall y[\exists z[(y, z) \in R \wedge (z, x) \in S] \rightarrow (y, x) \in (R \circ S)]$

C: therefore it follows: $(x, y) \in (S^{-1} \circ R^{-1}) \rightarrow (y, x) \in (R \circ S)$

	Statement A	Statement B	Statement C
PSYCOP	5	2	10
[Gentzen34]	3	3	9
Tutor			

Number of justifying proof steps for PSYCOP and Gentzen's NK.

Sample Analysis

A: $(x, y) \in (S^{-1} \circ R^{-1}) \Leftrightarrow \exists z[(z, x) \in S \wedge (y, z) \in R]$

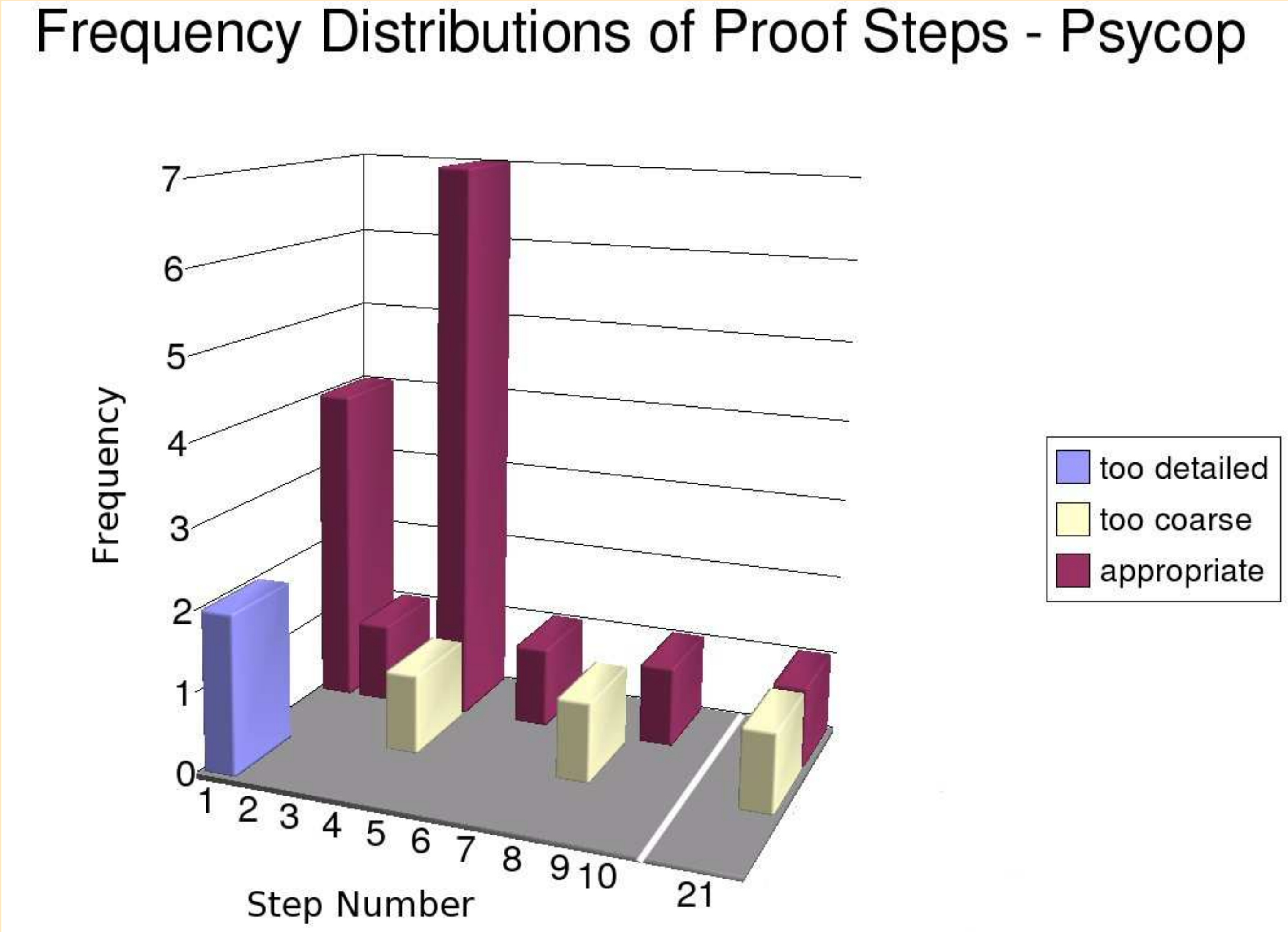
B: $\forall x \forall y[\exists z[(y, z) \in R \wedge (z, x) \in S] \rightarrow (y, x) \in (R \circ S)]$

C: therefore it follows: $(x, y) \in (S^{-1} \circ R^{-1}) \rightarrow (y, x) \in (R \circ S)$

	Statement A	Statement B	Statement C
PSYCOP	5	2	10
[Gentzen34]	3	3	9
Tutor	"too coarse-grained"	"appropriate"	"appropriate"

Number of justifying proof steps for PSYCOP and Gentzen's NK.

Discrimination?



Observations

- Sample of 20 statements: data not clear-cut (both for PSCYOP and Gentzen's NK).
- Dependence on natural language interpretation.
- Role of definitions?
- PSYCOP designed to explain reasoners untrained in formal logics.

Further Directions

1. Develop proof analysis mechanisms:
   - Enumeration & analysis of proof alternatives.
   - Develop & investigate complex evaluation hypotheses.
   - Develop & investigate cognitively “realistic” proof systems.
   - Relationship: granularity $\leftrightarrow$ relevance ?
2. Apply techniques and evaluate them empirically .

Questions?

Exercises from the experiment

Exercise W

Assume R and S are relations on an arbitrary set M. It holds:

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

Exercise A

Assume R, S and T are relations on an arbitrary set M. It holds:

$$(R \cup S) \circ T = (R \circ T) \cup (S \circ T)$$

Exercise B

Assume R, S and T are relations on a set M. It holds:

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

Exercises from the experiment

Exercise C

Assume R and S are relations on a set M. It holds:

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

Exercise E

Assume R is an asymmetric relation on a set M. Show: If E is not empty (i.e. $R \neq \emptyset$), then it holds:

$R \neq R^{-1}$