

# Mathematical Domain Reasoning Tasks in Natural Language Tutorial Dialog on Proofs

Christoph Benz Müller

Joint work with: SFB378 DIALOG Project



Computer Science & Comput. Ling.

Saarland University

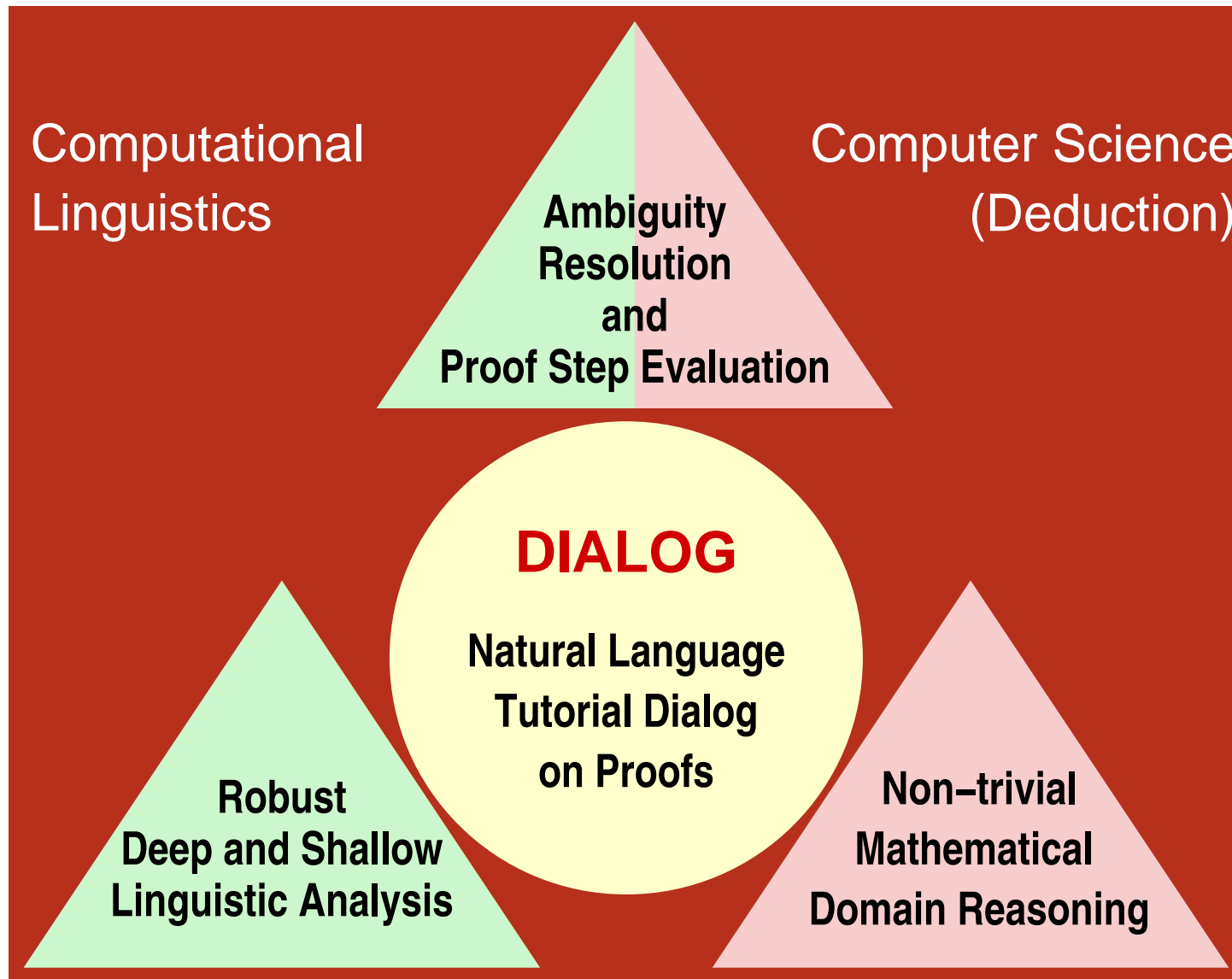
Saarbrücken, Germany

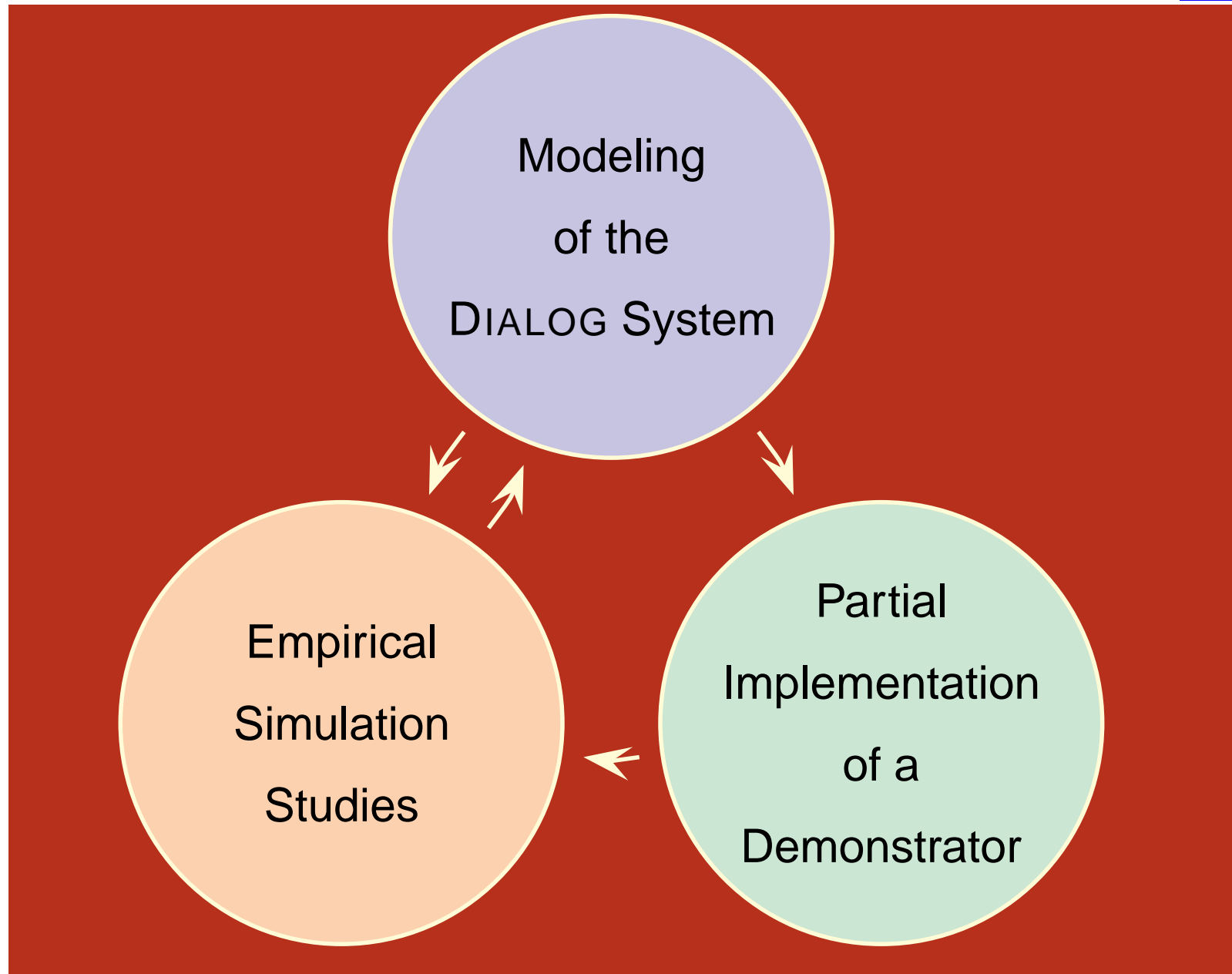
<http://www.ags.uni-sb.de/~chris/dialog/>

Dreamer Reunion, 1 August 2005, Edinburgh, UK

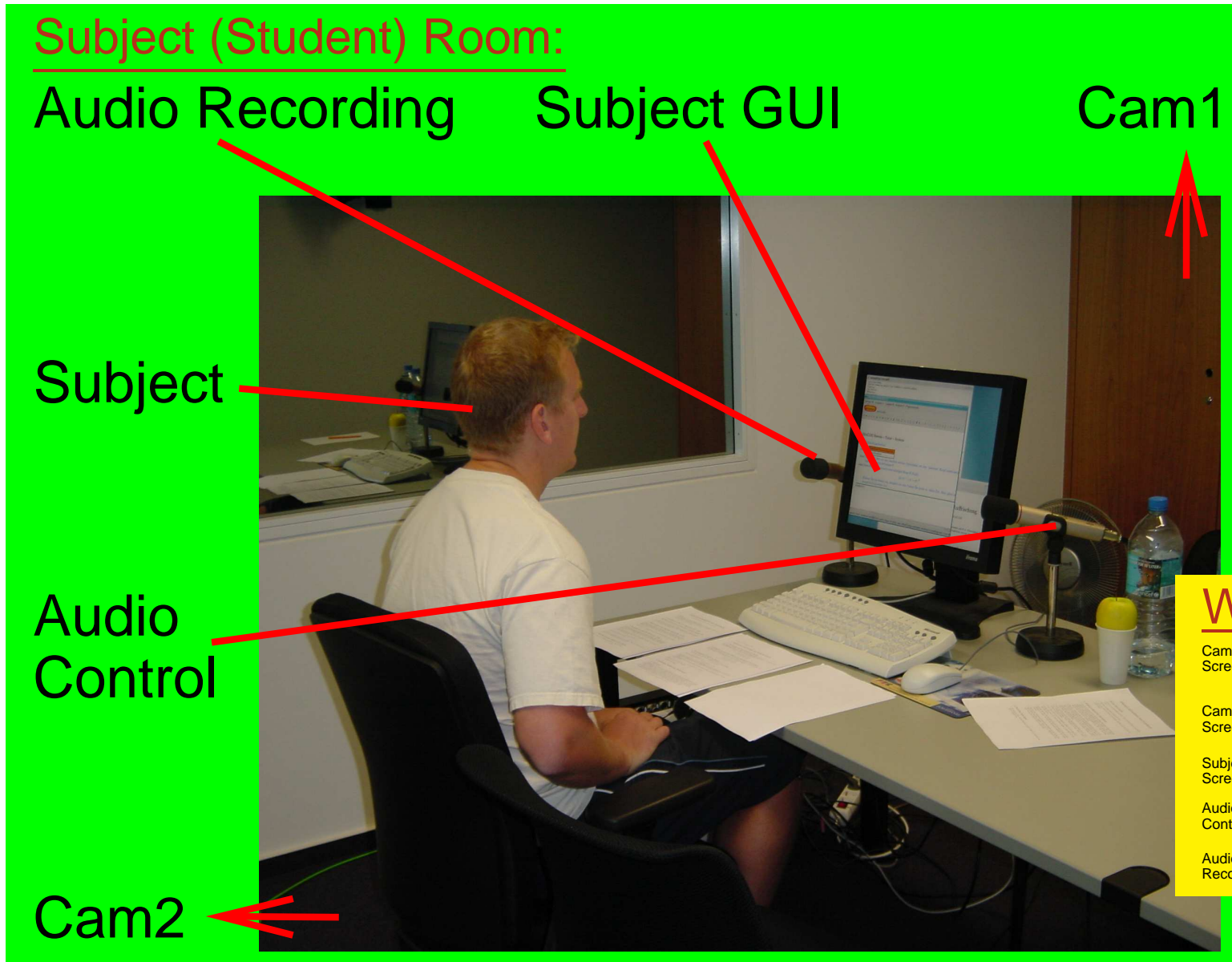
- $\Omega_{\text{MEGA}}-\lambda\text{Clam}$ : many research links, mutual research visits, friendships
- not in this talk
  - ▶  $\Omega_{\text{MEGA}}$ : basic research/development of an integrated mathematics assistance environment
- in this talk
  - ▶ DIALOG: NL based interaction with a mathematics assistance system
  - ▶ less well known to Dreamers
  - ▶ challenge between NL and AR
  - ▶ motivation for ‘abstract’-level reasoning: proof planning

# The DIALOG Project in the SFB 378





# WOZ-Experiment → Own Corpus



# WOZ-Experiment → Own Corpus

## Wizard (Tutor) Room:

Cam2  
Screen

Overall  
Control

Experi-  
menter

Wizard  
GUI

Wizard  
(Tutor)

Cam1  
Screen

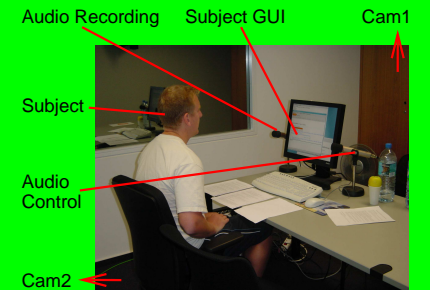
Subject  
Screen

Audio  
Control

Audio  
Recording



## Subject Room:





T1: Bitte zeigen Sie:  $K((A \cup B) \cap (C \cup D)) = (K(A) \cap K(B)) \cup (K(C) \cap K(D))!$

[Please show:  $K((A \cup B) \cap (C \cup D)) = (K(A) \cap K(B)) \cup (K(C) \cap K(D))!$ ]

S1: nach deMorgan-Regel-2 ist  $K((A \cup B) \cap (C \cup D)) = (K(A \cup B) \cup K(C \cup D))$ .

[by deMorgan-Rule-2  $K((A \cup B) \cap (C \cup D)) = (K(A \cup B) \cup K(C \cup D))$  holds.]

T2: Das ist richtig!

[This is correct!]

S2:  $K(A \cup B)$  ist laut deMorgan-1  $K(A) \cap K(B)$

[ $K(A \cup B)$  is  $K(A) \cap K(B)$  according to deMorgan-1]

T3: Das stimmt auch.

[That is also right.]

S3: und  $K(C \cup D)$  ist ebenfalls laut deMorgan-1  $K(C) \cap K(D)$

[and  $K(C \cup D)$  is also  $K(C) \cap K(D)$  according to deMorgan-1]

■ ■ ■

Get corpus: <http://www.ags.uni-sb.de/~chris/dialog/>

Total figures 1. exp.: 66 dialogs / av. 12 turns / 1115 sentences

## Perspective of Mathematical Domain Reasoning (MDR):

- Support for resolution of **Ambiguities** and **Underspecification**
- **Proof Step Evaluation**
  - ▶ **Soundness**: proof step verifiable by formal system?
  - ▶ **Granularity**: size/argumentative complexity of proof step?
  - ▶ **Relevance**: proof step needed/useful in achieving the goal?

Perspective of NL Analysis:

[... not in this talk ...]

Perspective of Dialog Management:

[... not in this talk ...]

Perspective of Tutoring Proofs:

[... not in this talk ...]

## Perspective of Mathematical Domain Reasoning (MDR):

- Support for resolution of **Ambiguity** and **Underspecification**
- **Proof Step Evaluation**
  - ▶ **Soundness:** proof verifiable by formal system?
  - ▶ **Granularity:** argumentative complexity of proof step?
  - ▶ **Relevance:** proof step needed/useful in achieving the goal?

**Logical vs Tutorial Dimension**

Perspective of NL Analysis:

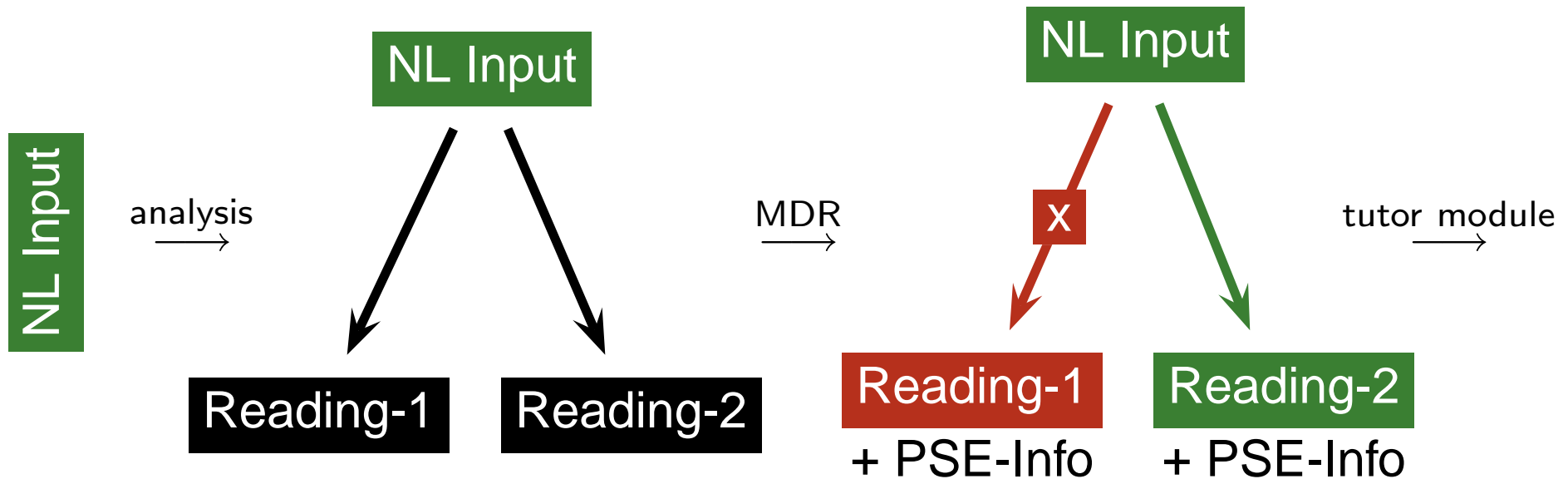
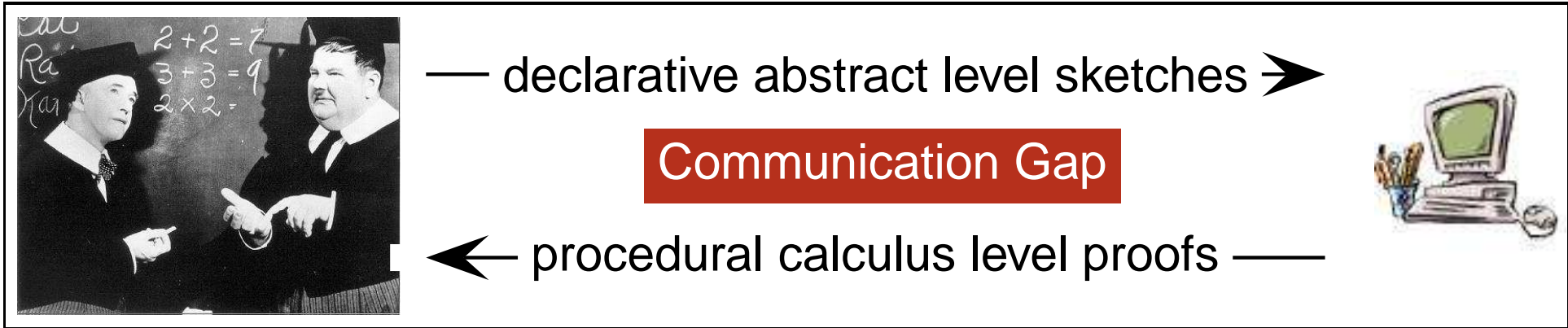
[... not in this talk ...]

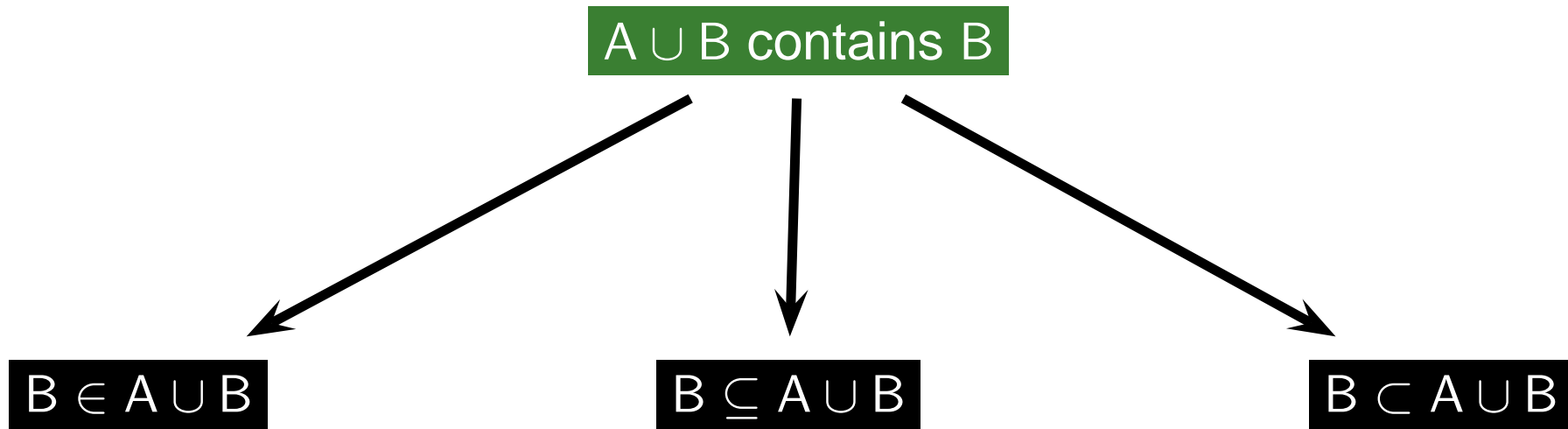
Perspective of Dialog Management:

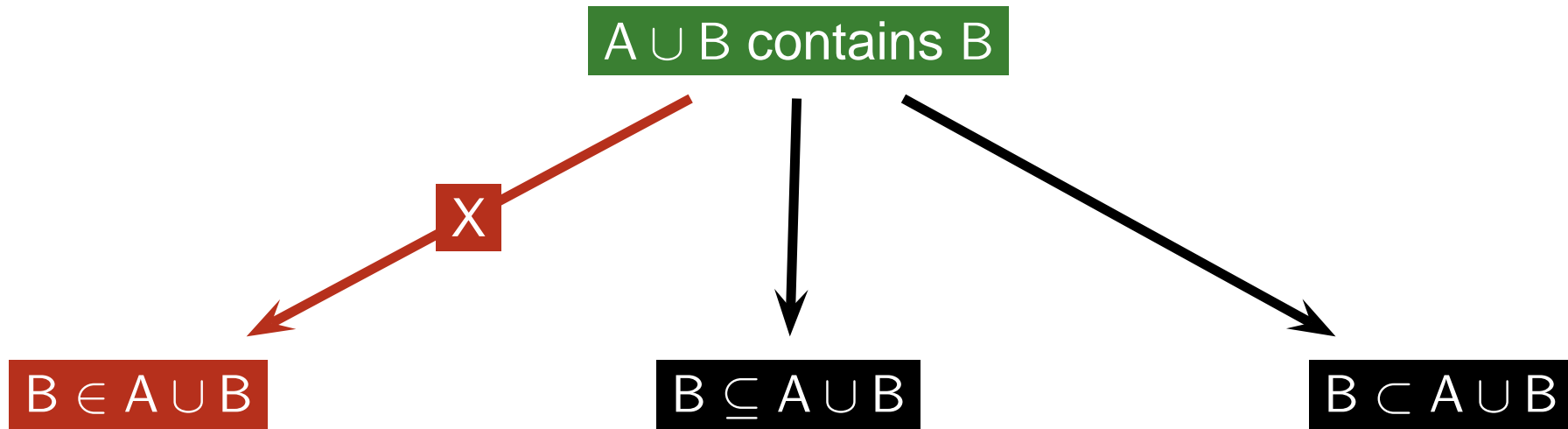
[... not in this talk ...]

Perspective of Tutoring Proofs:

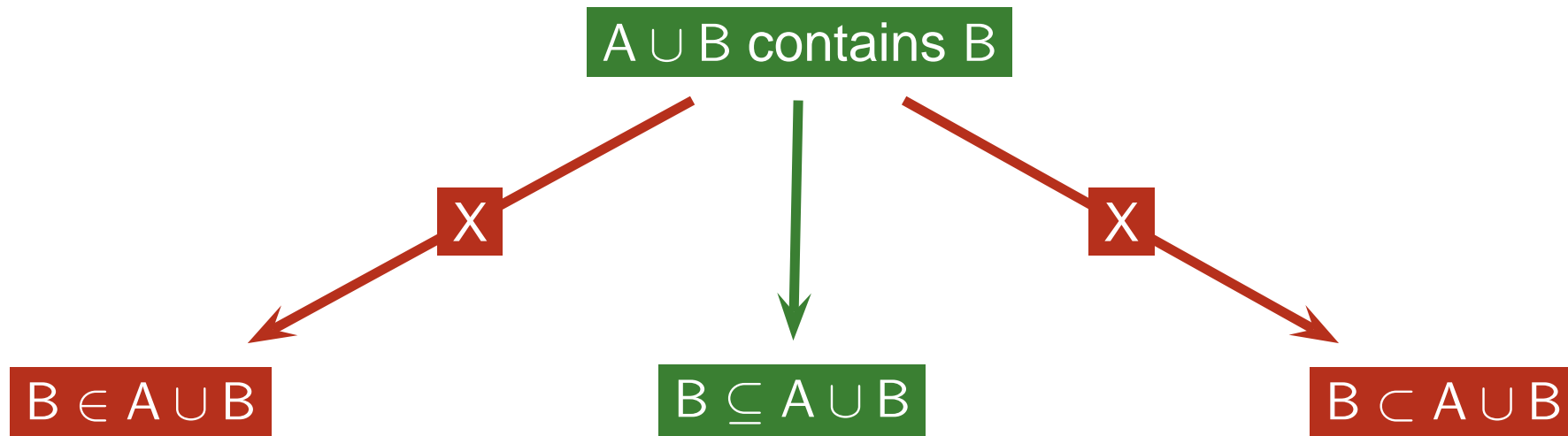
[... not in this talk ...]







type checking



theorem proving



$$\mathcal{P}((A \cup C) \cap (B \cup C)) = \mathcal{P}C \cup (A \cap B)$$

X

$$\mathcal{P}((A \cup C) \cap (B \cup C)) = \mathcal{P}(C) \cup (A \cap B)$$

$$\mathcal{P}((A \cup C) \cap (B \cup C)) = \mathcal{P}(C \cup (A \cap B))$$

type checking



$$\kappa((A \cup C) \cap (B \cup C)) = \kappa C \cup (A \cap B)$$

X

$$\kappa((A \cup C) \cap (B \cup C)) = \kappa(C) \cup (A \cap B)$$

$$\kappa((A \cup C) \cap (B \cup C)) = \kappa(C \cup (A \cap B))$$

theorem proving

(DM-1) ...

(DM-2) ...

?

(G) ...

Given: (DM-1)  $\overline{X \cup Y} = \bar{X} \cap \bar{Y}$

(DM-2)  $\overline{X \cap Y} = \bar{X} \cup \bar{Y}$

**Task:** Please show  $\overline{(A \cup B) \cap (C \cup D)} = (\bar{A} \cap \bar{B}) \cup (\bar{C} \cap \bar{D})$

New: By deMorgan  $\overline{(A \cup B) \cap (C \cup D)} = \overline{(A \cup B)} \cup \overline{(C \cup D)}$ .

# Proof Step Evaluation

(DM-1) ...  
(DM-2) ...  
?  
(G) ...

Given: (DM-1)  $\overline{X \cup Y} = \overline{X} \cap \overline{Y}$

(DM-2)  $\overline{X \cap Y} = \overline{X} \cup \overline{Y}$

**Task:** Please show  $\overline{(A \cup B) \cap (C \cup D)} = (\overline{A} \cap \overline{B}) \cup (\overline{C} \cap \overline{D})$

New: By deMorgan  $\overline{(A \cup B) \cap (C \cup D)} = \overline{(A \cup B)} \cup \overline{(C \cup D)}$ .

(DM-1) ...  
(DM-2) ...  
(New) ... ↓  
?  
(G) ...

Soundness: yes

Granularity: 1x(DM-2)

Relevance: yes

(DM-1) ...  
(DM-2) ...  
?  
(New) ... ↑  
(G) ...

Soundness: yes

Granularity: 2x(DM-1)

Relevance: yes

# Proof Step Evaluation: How?

Discourse:

- (1)  $A \wedge B$
- (2)  $A \Rightarrow C$
- (3)  $C \Rightarrow D$
- (4)  $F \Rightarrow B$
- ?
- (G)  $D \vee E$

New:

We show E.



- (1) ...
- (2) ...
- (3) ...
- (4) ...
- ?
- (G') E
- (G) ...

PSE:

Soundness

Granularity

Relevance

# Proof Step Evaluation: How?

Discourse:

(1)  $A \wedge B$   
(2)  $A \Rightarrow C$   
(3)  $C \Rightarrow D$   
(4)  $F \Rightarrow B$   
?  
**(G)  $D \vee E$**

New:

We show E.



(1) ...  
(2) ...  
(3) ...  
(4) ...  
?  
**(G') E**  
(G) ...

PSE:

Soundness

- $(G') \vdash^? (G)$
- any proof

Granularity

Relevance

# Proof Step Evaluation: How?

## Discourse:

(1)  $A \wedge B$   
(2)  $A \Rightarrow C$   
(3)  $C \Rightarrow D$   
(4)  $F \Rightarrow B$   
?  
 $(G) D \vee E$

## New:

We show E.



(1) ...  
(2) ...  
(3) ...  
(4) ...  
?  
 $(G') E$   
 $(G) \dots$

## PSE:

### Soundness

- $(G') \vdash^? (G)$
- any proof

### Granularity

- $\text{complexity}((G') \vdash^? (G))$
- cognitively adequate proofs

### Relevance

# Proof Step Evaluation: How?

Discourse:

(1)  $A \wedge B$   
(2)  $A \Rightarrow C$   
(3)  $C \Rightarrow D$   
(4)  $F \Rightarrow B$   
?  
**(G)  $D \vee E$**

New:

We show E.



(1) ...  
(2) ...  
(3) ...  
(4) ...  
?  
**(G') E**  
**(G) ...**

PSE:

**Soundness**

- $(G') \vdash^? (G)$
- any proof

**Granularity**

- $\text{complexity}((G') \vdash^? (G))$
- cognitively adequate proofs

**Relevance**

- $(1), (2), (3), (4) \vdash^? (G')$
- detours?, shorter proofs?

## Granularity and Relevance call for

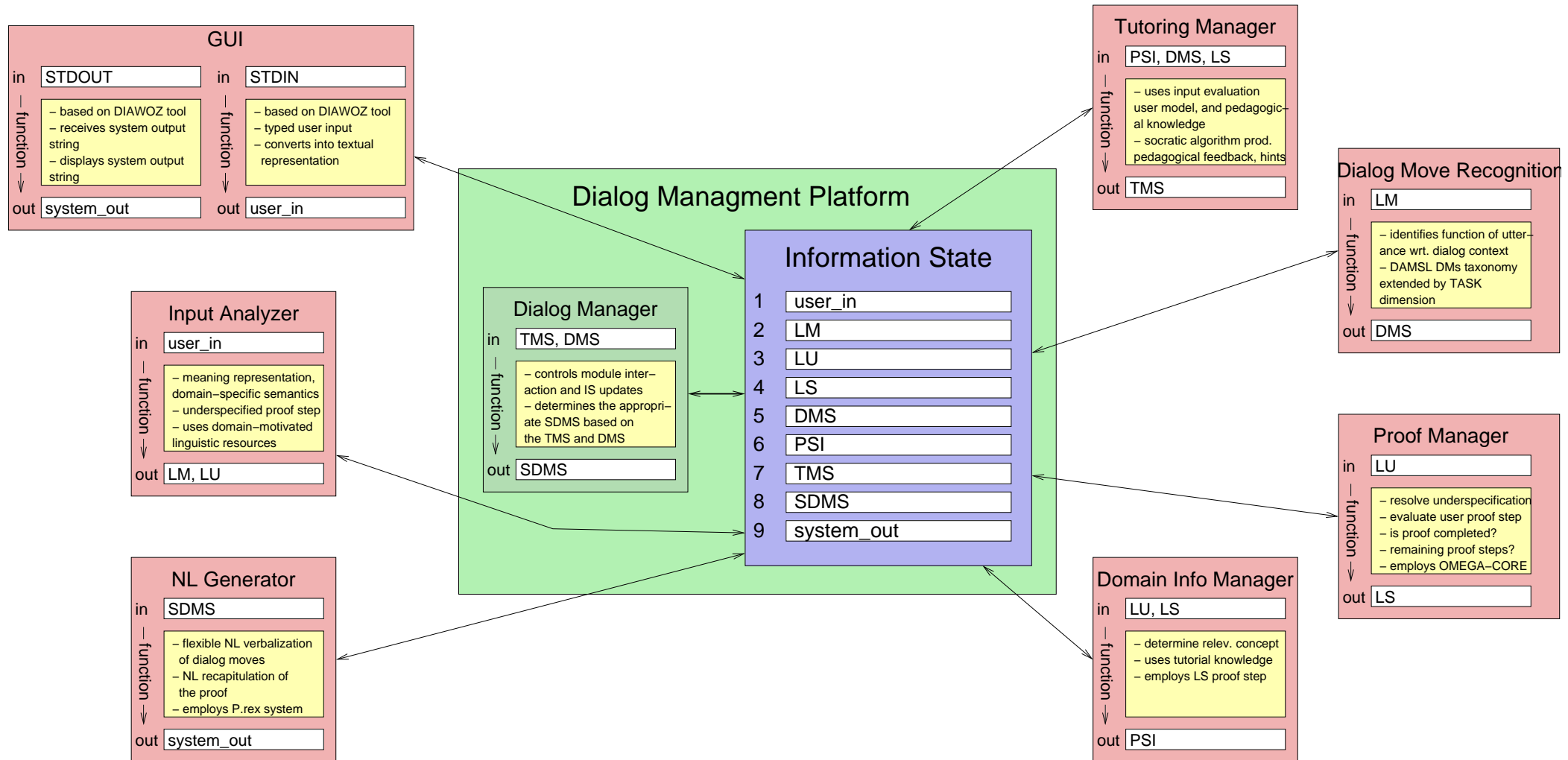
cognitively adequate abstract level proofs

+

enumeration of (some) proof alternatives

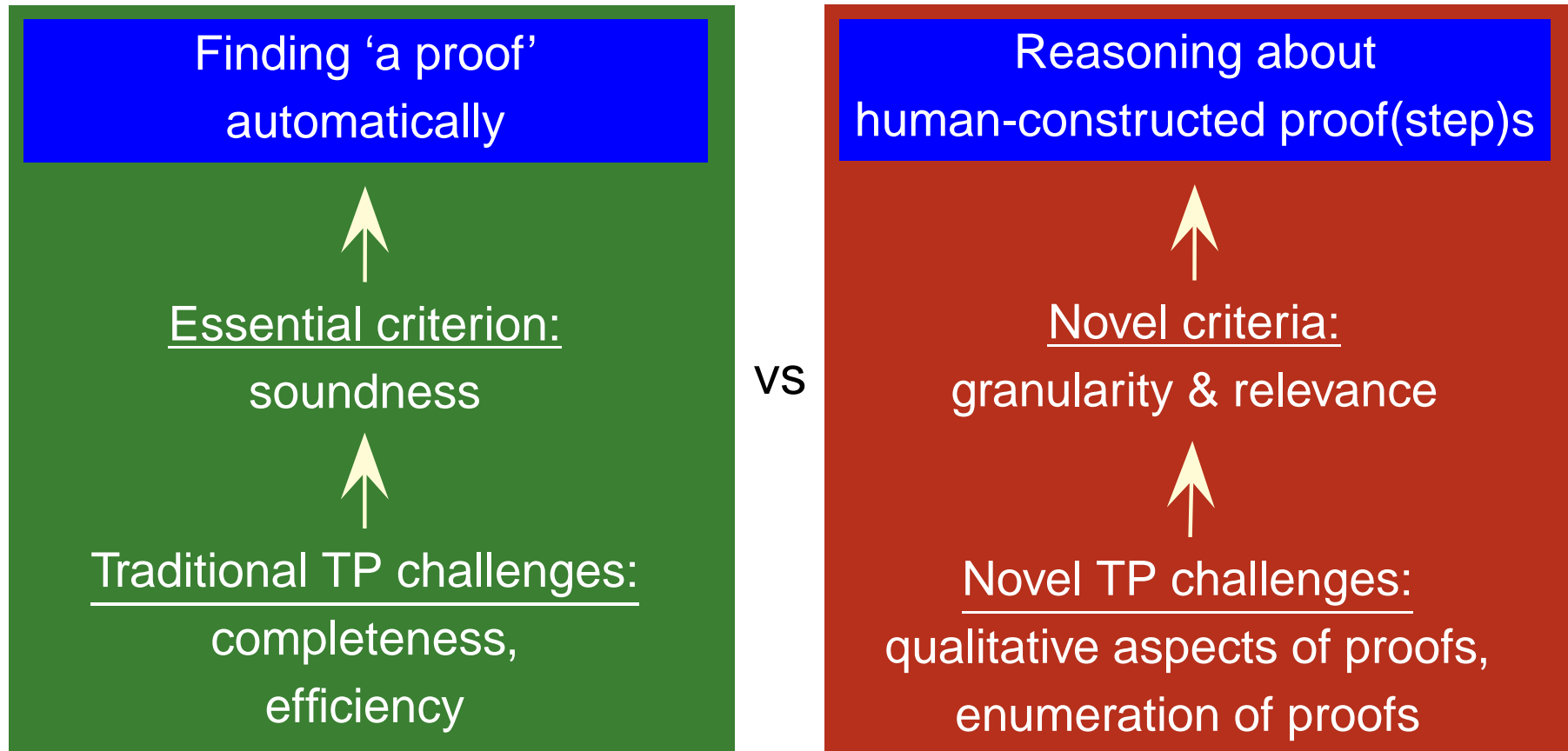
- One candidate: knowledge based proof planning [Bundy88]
- Original motivation: widen range of automatable maths
- New motivation: support for proof step evaluation

# Implementation: DIALOG Demonstrator





- Motivation: [Moore93] Flexible tutorial NL dialog supports active learning
- Closest related: [Zinn04] analyzes well structured text-book proofs; lots of interesting ongoing work
- NL analysis: shallow techniques and keyword spotting probably not suitable
- MDR: Comparison against 'golden standard solutions' [GreaserEtAl00] not suitable
- Dialog modeling: Autotutor [PersonEtAl00], Geometry Tutor [MatsudaVanLehn03], Trindi and Siridus [TraumLarsson03], Beetle [Zinn03]



- Lots of ongoing work in all corners of the DIALOG Project