Presenting Proofs with Adapted Granularity

Marvin Schiller[†] and Christoph Benzmüller[⋆]

[†]Saarland University, Germany
[†]DFKI GmbH, Germany
[⋆]Articulate Software, Angwin, CA, USA

KI 2009 – 32nd Annual Conference on AI,
September 15th – 18th, 2009

Studienstiftung des deutschen Volkes

UNIVERSITÄT DES SAARLANDES

$\boxed{1}$ *"Let* x *be an element of* $A \cap (B \cup C)$, $\boxed{2}$ *then* $x \in A$ *and* $x \in B \cup C$. $\boxed{3}$ *This means that* $x \in A$, *and either* $x \in B$ *or* $x \in C$. $\boxed{4}$ *Hence we either have (i)* $x \in A$ *and* $x \in B$, *or we have (ii)* $x \in A$ *and* $x \in C$. $\boxed{5}$ *Therefore, either* $x \in A \cap B$ *or* $x \in A \cap C$, *so* $\boxed{6}$ $x \in (A \cap B) \cup (A \cap C)$. $\boxed{7}$ *This shows that* $A \cap (B \cup C)$ *is a subset of* $(A \cap B) \cup (A \cap C)$. $\boxed{8}$ *Conversely, let* y *be an element of* $(A \cap B) \cup (A \cap C)$. $\boxed{9}$ *Then, either (iii)* $y \in A \cap B$, *or (iv)* $y \in A \cap C$. $\boxed{10}$ *It follows that* $y \in A$, *and either* $y \in B$ *or* $y \in C$. $\boxed{11}$ *Therefore,* $y \in A$ *and* $y \in B \cup C$ $\boxed{12}$ *so that* $y \in A \cap (B \cup C)$. $\boxed{13}$ *Hence* $(A \cap B) \cup (A \cap C)$ *is a subset of* $A \cap (B \cup C)$. $\boxed{14}$ *In view of Definition 1.1.1, we conclude that the sets* $A \cap (B \cup C)$ *and* $(A \cap B) \cup (A \cap C)$ *are equal."*

Proof of $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, reproduced from Bartle/Sherbert: Introduction to Real Analysis, 1982.

Proof Step Size in Mathematics Instruction

$\boxed{9}$ Then, either (iii) $y \in A \cap B$, or (iv) $y \in A \cap C$.
$\boxed{10a}$ $y \in A \wedge y \in B$, or $y \in A \cap C$ (Def. $\cap$)
$\boxed{10b}$ $y \in A \wedge y \in B$, or $y \in A \wedge y \in C$ (Def. $\cap$)
$\boxed{10c}$ It follows that $y \in A$, and either $y \in B$ or $y \in C$. (Distr.)

then $x \in A$ and $x \in B \cup C$.
$x \in B$ or $x \in C$. $\boxed{4}$ Hence
, or we have (ii) $x \in A$
$A \cap B$ or $x \in A \cap C$, so
$\boxed{6}$ $x \in (A \cap B) \cup (A \cap C)$. $\boxed{7}$ This shows that $A \cap (B \cup C)$ is a subset of $(A \cap B) \cup (A \cap C)$. $\boxed{8}$ Conversely, let y be an element of $(A \cap B) \cup (A \cap C)$. $\boxed{9}$ Then, either (iii) $y \in A \cap B$, or (iv) $y \in A \cap C$. $\boxed{10}$ It follows that $y \in A$, and either $y \in B$ or $y \in C$. $\boxed{11}$ Therefore, $y \in A$ and $y \in B \cup C$ $\boxed{12}$ so that $y \in A \cap (B \cup C)$. $\boxed{13}$ Hence $(A \cap B) \cup (A \cap C)$ is a subset of $A \cap (B \cup C)$. $\boxed{14}$ In view of Definition 1.1.1, we conclude that the sets $A \cap (B \cup C)$ and $(A \cap B) \cup (A \cap C)$ are equal."

Proof of $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, reproduced from Bartle/Sherbert: Introduction to Real Analysis, 1982.

Overview

1. Framework for Granularity-Adaptive Proof Presentation
   - Approach
   - Case Study

2. An Empirical Study on Granularity

3. Conclusion

Granularity-Adaptive Proof Presentation Framework

Goal: bridge granularity gap between theorem proving systems/mathematical practice

- parameterized over 'granularity policy'
- adaptive (e.g. to user's/learner's knowledge)
- implementation (here, using MAS ΩMEGA) & experiments

Previous Work

- HiProofs [Denney et al. 2006], proof presentation in Isabelle [Simons, 1997], Theorema [Buchberger et al., 1997], Coq [Coscoy, Kahn, Thery, 1995], etc.
- hierarchical proof planning, proof explainer P.rex [Fiedler 2001]
- particular granularity levels, e.g. *what-you-need-is-what-you-stated* [Autexier/Fiedler 2006]

Granularity as a Classification Problem

Approach

- Granularity as a classification problem; simple/composite steps can be *appropriate*, *too big* or *too small*
- Identify properties of proof steps that can make them appropriate/inappropriate (granularity criteria)

E.g.

$\boxed{9}$ Then, either (iii) $y \in A \cap B$, or (iv) $y \in A \cap C$.
$\boxed{10}$ It follows that $y \in A$, and either $y \in B$ or $y \in C$.

(i) two concepts involved: def. of $\cap$, distributivity,
(ii) total number of concept applications: three,
(iii) involved concepts have been previously applied,
(iv) all manipulations apply to a common part in $\boxed{9}$,
(v) the names of the applied concepts are not explicitly mentioned
(vi) inferences belong to naive set theory and propositional logic.

- Model decision which steps are skipped/combined via classic expert system

Algorithm

Main loop (input: proof tree)

(i) consider (iteratively larger) proof segments
(ii) analyze granularity-relevant properties of each composite step
(iii) judge each composite step via granularity classifier (assign labels "too big" / "too small" / "appropriate"), with and without mentioning concept name(s)
(iv) select composite steps of "appropriate" size
(v) pretty print & simple NL (pattern-based)

'Granularity policies' expressed as classifiers (e.g. rule sets/decision trees), e.g.

* concepts $\in\{0,1\} \wedge$ eq-Defn$=0 \wedge$ verb$=$true $\Rightarrow$ too-small
* concepts $\in\{2,3,4\} \wedge \cup$-Defn $\in\{1,2,3\} \Rightarrow$ too-big

$\vdots$

* $\_ \Rightarrow$ appropriate

Classifiers

... model different levels of proof granularity.

Classifiers can be

- hand-authored
- fitted to sample proofs (using machine learning)
- learned (via ML) from annotated proof samples

Granularity Criteria

- mastered/unmastered concepts (w.r.t. simple overlay student model)
- are hypotheses/subgoals introduced?
- direction (forward/backward)
- explicitness (explanation)
- etc...

Example

$$x \in \mathbf{S} \vdash x \in \mathbf{S}$$
$\text{Def}\cup\ (8)$
$$(x\in(A\cap B)\vee x\in(A\cap C)) \vdash x\in\mathbf{S}$$
$\text{Def}\cap\ (7)$
$$(x\in(A\cap B)\vee x\in A\wedge x\in C) \vdash x\in\mathbf{S}$$
$\text{Def}\cap\ (6)$
$$(x\in A\wedge x\in B\vee x\in A\wedge x\in C)\vdash x\in\mathbf{S}$$
$\text{DISTR}\ (5)$
$$(x\in A\wedge(x\in B \vee x\in C)) \vdash x\in\mathbf{S}$$
$\text{Def}\cup\ (4)$
$$(x\in A\wedge x\in(B\cup C)) \vdash x\in\mathbf{S}$$
$\text{Def}\cap\ (3)$
$$(x\in(A\cap(B\cup C))) \vdash x\in\mathbf{S}$$
$\text{Def}\subseteq\ (2)$
$$\vdash (A\cap(B\cup C))\subseteq\mathbf{S}$$

$$y\in\mathbf{T} \vdash y\in\mathbf{T}$$
$\text{Def}\cap\ (15)$
$$(y\in A \wedge y\in(B\cup C)) \vdash y\in\mathbf{T}$$
$\text{Def}\cup\ (14)$
$$(y\in A\wedge(y\in B\vee y\in C))\vdash y\in\mathbf{T}$$
$\text{DISTR}\ (13)$
$$(y\in A\wedge y\in B\vee y\in A\wedge y\in C)\vdash y\in\mathbf{T}$$
$\text{Def}\cap\ (12)$
$$(y\in A\wedge y\in B\vee y\in(A\cap C)) \vdash y\in\mathbf{T}$$
$\text{Def}\cap\ (11)$
$$(y\in(A\cap B) \vee y \in (A\cap C)) \vdash y\in\mathbf{T}$$
$\text{Def}\cup\ (10)$
$$(y\in((A\cap B)\cup(A\cap C))) \vdash y\in\mathbf{T}$$
$\text{Def}\subseteq\ (9)$
$$\vdash ((A\cap B)\cup(A\cap C)) \subseteq \mathbf{T}$$

$\text{Def EQ}\ (1)$
$$\vdash \underbrace{(A\cap(B\cup C))}_{\mathbf{T}} = \underbrace{((A\cap B)\cup(A\cap C))}_{\mathbf{S}}$$

: appropriate

: appropriate with explanation

: too small (skipped)

Example

$$x \in \mathbf{S} \vdash x \in \mathbf{S}$$
$$\text{Def}\cup\ (8)$$
$$\boxed{(x\in(A\cap B)\vee x\in(A\cap C)) \vdash x\in\mathbf{S}}$$
$$\text{Def}\cap\ (7)$$
$$\boxed{(x\in(A\cap B)\vee x\in A\wedge x\in C)\vdash x\in\mathbf{S}}$$
$$\text{Def}\cap\ (6)$$
$$\boxed{(x\in A\wedge x\in B\vee x\in A\wedge x\in C)\vdash x\in\mathbf{S}}$$
$$\text{DISTR}\ (5)$$
$$\boxed{(x\in A\wedge(x\in B\vee x\in C))\vdash x\in\mathbf{S}}$$
$$\text{Def}\cup\ (4)$$
$$\boxed{(x\in A\wedge x\in(B\cup C))\vdash x\in\mathbf{S}}$$
$$\text{Def}\cap\ (3)$$
$$\boxed{(x\in(A\cap(B\cup C)))\vdash x\in\mathbf{S}}$$
$$\text{Def}\subseteq\ (2)$$
$$\boxed{\vdash (A\cap(B\cup C))\subseteq\mathbf{S}}$$

$$y\in\mathbf{T}\vdash y\in\mathbf{T}$$
$$\text{Def}\cap\ (15)$$
$$(y\in A\wedge y\in(B\cup C))\vdash y\in\mathbf{T}$$
$$\text{Def}\cup\ (14)$$
$$(y\in A\wedge(y\in B\vee y\in C))\vdash y\in\mathbf{T}$$
$$\text{DISTR}\ (13)$$
$$(y\in A\wedge y\in B\vee y\in A\wedge y\in C)\vdash y\in\mathbf{T}$$
$$\text{Def}\cap\ (12)$$
$$\boxed{(y\in A\wedge y\in B\vee y\in(A\cap C))\vdash y\in\mathbf{T}}$$
$$\text{Def}\cap\ (11)$$
$$\boxed{(y\in(A\cap B)\vee y\in(A\cap C))\vdash y\in\mathbf{T}}$$
$$\text{Def}\cup\ (10)$$
$$\boxed{(y\in((A\cap B)\cup(A\cap C)))\vdash y\in\mathbf{T}}$$
$$\text{Def}\subseteq\ (9)$$
$$\boxed{\vdash((A\cap B)\cup(A\cap C))\subseteq\mathbf{T}}$$

$$\text{Def Eq}\ (1)$$
$$\boxed{\vdash \underbrace{(A\cap(B\cup C))}_{\mathbf{T}} = \underbrace{((A\cap B)\cup(A\cap C))}_{\mathbf{S}}}$$

$$\frac{\boxed{(y\in A\wedge y\in B\vee y\in(A\cap C))\vdash y\in\mathbf{T}}}{\boxed{(y\in(A\cap B)\vee y\in(A\cap C))\vdash y\in\mathbf{T}}}\ \text{Def}\cap$$

Properties: concepts:1, total: 1, mastered: 1, etc...

Classification result: "too small" (w. and w./o. explanation)

Example

$$
\begin{array}{rc}
 & x \in \mathbf{S} \vdash x \in \mathbf{S} \\ \hline
\text{Def}\cup\ (8) & (x\in(A\cap B)\vee x\in(A\cap C)) \vdash x\in\mathbf{S} \\ \hline
\text{Def}\cap\ (7) & (x\in(A\cap B)\vee x\in A\wedge x\in C)\vdash x\in\mathbf{S} \\ \hline
\text{Def}\cap\ (6) & (x\in A\wedge x\in B\vee x\in A\wedge x\in C)\vdash x\in\mathbf{S} \\ \hline
\text{DISTR}\ (5) & (x\in A\wedge(x\in B\vee x\in C))\vdash x\in\mathbf{S} \\ \hline
\text{Def}\cup\ (4) & (x\in A\wedge x\in(B\cup C))\vdash x\in\mathbf{S} \\ \hline
\text{Def}\cap\ (3) & (x\in(A\cap(B\cup C)))\vdash x\in\mathbf{S} \\ \hline
\text{Def}\subseteq\ (2) & \vdash (A\cap(B\cup C))\subseteq\mathbf{S}
\end{array}
\qquad
\begin{array}{cl}
y\in\mathbf{T}\vdash y\in\mathbf{T} & \\ \hline
(y\in A\wedge y\in(B\cup C))\vdash y\in\mathbf{T} & \text{Def}\cap\ (15) \\ \hline
(y\in A\wedge(y\in B\vee y\in C))\vdash y\in\mathbf{T} & \text{Def}\cup\ (14) \\ \hline
(y\in A\wedge y\in B\vee y\in A\wedge y\in C)\vdash y\in\mathbf{T} & \text{DISTR}\ (13) \\ \hline
(y\in A\wedge y\in B\vee y\in(A\cap C))\vdash y\in\mathbf{T} & \text{Def}\cap\ (12) \\ \hline
(y\in(A\cap B)\vee y\in(A\cap C))\vdash y\in\mathbf{T} & \text{Def}\cap\ (11) \\ \hline
(y\in((A\cap B)\cup(A\cap C)))\vdash y\in\mathbf{T} & \text{Def}\cup\ (10) \\ \hline
\vdash((A\cap B)\cup(A\cap C))\subseteq\mathbf{T} & \text{Def}\subseteq\ (9)
\end{array}
$$

$$
\text{Def Eq}\ (1)\quad \vdash \underbrace{(A\cap(B\cup C))}_{\mathbf{T}} = \underbrace{((A\cap B)\cup(A\cap C))}_{\mathbf{S}}
$$

$$
\begin{array}{cl}
(y\in A\wedge y\in B\vee y\in A\wedge y\in C)\vdash y\in\mathbf{T} & \\ \hline
(y\in A\wedge y\in B\vee y\in(A\cap C))\vdash y\in\mathbf{T} & \text{Def}\cap \\ \hline
(y\in(A\cap B)\vee y\in(A\cap C))\vdash y\in\mathbf{T} & \text{Def}\cap
\end{array}
$$

Properties:
concepts:1,
total: 2,
mastered: 1,
etc...

Classification result: "too small" (w. and w./o. explanation)

Example

$$
\begin{array}{rc}
 & \boxed{x\in \mathbf{S} \vdash x\in \mathbf{S}} \\ \hline
\text{Def}\cup\ (8) & (x\in(A\cap B)\vee x\in(A\cap C)) \vdash x\in \mathbf{S} \\ \hline
\text{Def}\cap\ (7) & (x\in(A\cap B)\vee x\in A\wedge x\in C) \vdash x\in \mathbf{S} \\ \hline
\text{Def}\cap\ (6) & (x\in A\wedge x\in B\vee x\in A\wedge x\in C)\vdash x\in \mathbf{S} \\ \hline
\text{DISTR}\ (5) & (x\in A\wedge(x\in B\vee x\in C)) \vdash x\in \mathbf{S} \\ \hline
\text{Def}\cup\ (4) & (x\in A\wedge x\in(B\cup C)) \vdash x\in \mathbf{S} \\ \hline
\text{Def}\cap\ (3) & (x\in(A\cap(B\cup C))) \vdash x\in \mathbf{S} \\ \hline
\text{Def}\subseteq\ (2) & \vdash (A\cap(B\cup C))\subseteq \mathbf{S}
\end{array}
\qquad
\begin{array}{cl}
y\in \mathbf{T} \vdash y\in \mathbf{T} & \\ \hline
(y\in A\wedge y\in(B\cup C)) \vdash y\in \mathbf{T} & \text{Def}\cap\ (15) \\ \hline
\boxed{(y\in A\wedge(y\in B\vee y\in C))\vdash y\in \mathbf{T}} & \text{Def}\cup\ (14) \\ \hline
(y\in A\wedge y\in B\vee y\in A\wedge y\in C)\vdash y\in \mathbf{T} & \text{DISTR}\ (13) \\ \hline
(y\in A\wedge y\in B\vee y\in(A\cap C)) \vdash y\in \mathbf{T} & \text{Def}\cap\ (12) \\ \hline
(y\in(A\cap B)\vee y\in(A\cap C)) \vdash y\in \mathbf{T} & \text{Def}\cap\ (11) \\ \hline
(y\in((A\cap B)\cup(A\cap C))) \vdash y\in \mathbf{T} & \text{Def}\cup\ (10) \\ \hline
\vdash ((A\cap B)\cup(A\cap C)) \subseteq \mathbf{T} & \text{Def}\subseteq\ (9)
\end{array}
$$

$$
\text{Def Eq}\ (1)\quad \vdash \underbrace{(A\cap(B\cup C))}_{\mathbf{T}} = \underbrace{((A\cap B)\cup(A\cap C))}_{\mathbf{S}}
$$

$$
\begin{array}{cl}
\boxed{(y\in A\wedge(y\in B\vee y\in C))\vdash y\in \mathbf{T}} & \\ \hline
(y\in A\wedge y\in B\vee y\in A\wedge y\in C)\vdash y\in \mathbf{T} & \text{DISTR} \\ \hline
(y\in A\wedge y\in B\vee y\in(A\cap C)) \vdash y\in \mathbf{T} & \text{Def}\cap \\ \hline
(y\in(A\cap B)\vee y\in(A\cap C)) \vdash y\in \mathbf{T} & \text{Def}\cap
\end{array}
$$

Properties: concepts:2, total: 3, mastered: 2, etc...

Classification result: "appropriate" (w./o. explanation)

Example

$$\dfrac{x \in \mathbf{S} \vdash x \in \mathbf{S}}{(x\in(A\cap B)\vee x\in(A\cap C)) \vdash x\in\mathbf{S}}\ \textsc{Def}\cup\ (8)$$

$$\dfrac{(x\in(A\cap B)\vee x\in(A\cap C)) \vdash x\in\mathbf{S}}{(x\in(A\cap B)\vee x\in A\wedge x\in C)\vdash x\in\mathbf{S}}\ \textsc{Def}\cap\ (7)$$

$$\dfrac{(x\in(A\cap B)\vee x\in A\wedge x\in C)\vdash x\in\mathbf{S}}{(x\in A\wedge x\in B\vee x\in A\wedge x\in C)\vdash x\in\mathbf{S}}\ \textsc{Def}\cap\ (6)$$

$$\dfrac{(x\in A\wedge x\in B\vee x\in A\wedge x\in C)\vdash x\in\mathbf{S}}{(x\in A\wedge(x\in B \vee x\in C)) \vdash x\in\mathbf{S}}\ \text{DISTR}\ (5)$$

$$\dfrac{(x\in A\wedge(x\in B \vee x\in C)) \vdash x\in\mathbf{S}}{(x\in A\wedge x\in(B\cup C)) \vdash x\in\mathbf{S}}\ \textsc{Def}\cup\ (4)$$

$$\dfrac{(x\in A\wedge x\in(B\cup C)) \vdash x\in\mathbf{S}}{(x\in(A\cap(B\cup C))) \vdash x\in\mathbf{S}}\ \textsc{Def}\cap\ (3)$$

$$\dfrac{(x\in(A\cap(B\cup C))) \vdash x\in\mathbf{S}}{\vdash (A\cap(B\cup C))\subseteq\mathbf{S}}\ \textsc{Def}\subseteq\ (2)$$

$$\dfrac{y\in\mathbf{T} \vdash y\in\mathbf{T}}{(y\in A \wedge y\in(B\cup C)) \vdash y\in\mathbf{T}}\ \textsc{Def}\cap\ (15)$$

$$\dfrac{(y\in A \wedge y\in(B\cup C)) \vdash y\in\mathbf{T}}{(y\in A\wedge(y\in B\vee y\in C))\vdash y\in\mathbf{T}}\ \textsc{Def}\cup\ (14)$$

$$\dfrac{(y\in A\wedge(y\in B\vee y\in C))\vdash y\in\mathbf{T}}{(y\in A\wedge y\in B\vee y\in A\wedge y\in C)\vdash y\in\mathbf{T}}\ \text{DISTR}\ (13)$$

$$\dfrac{(y\in A\wedge y\in B\vee y\in A\wedge y\in C)\vdash y\in\mathbf{T}}{(y\in A\wedge y\in B\vee y\in(A\cap C)) \vdash y\in\mathbf{T}}\ \textsc{Def}\cap\ (12)$$

$$\dfrac{(y\in A\wedge y\in B\vee y\in(A\cap C)) \vdash y\in\mathbf{T}}{(y\in(A\cap B) \vee y \in (A\cap C)) \vdash y\in\mathbf{T}}\ \textsc{Def}\cap\ (11)$$

$$\dfrac{(y\in(A\cap B) \vee y \in (A\cap C)) \vdash y\in\mathbf{T}}{(y\in((A\cap B)\cup(A\cap C))) \vdash y\in\mathbf{T}}\ \textsc{Def}\cup\ (10)$$

$$\dfrac{(y\in((A\cap B)\cup(A\cap C))) \vdash y\in\mathbf{T}}{\vdash ((A\cap B)\cup(A\cap C)) \subseteq \mathbf{T}}\ \textsc{Def}\subseteq\ (9)$$

$$\dfrac{\vdash (A\cap(B\cup C))\subseteq\mathbf{S} \qquad \vdash ((A\cap B)\cup(A\cap C)) \subseteq \mathbf{T}}{\vdash \underbrace{(A\cap(B\cup C))}_{\mathbf{T}} = \underbrace{((A\cap B)\cup(A\cap C))}_{\mathbf{S}}}\ \textsc{Def eq}\ (1)$$

Output for the Running Example

1) 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
2) We assume $x \in A \cap B \cup C$ and show $x \in (A \cap B) \cup (A \cap C)$
3) Therefore, $x \in A \wedge x \in B \cup C$
4) Therefore, $x \in A \wedge (x \in B \vee x \in C)$
5) Therefore, $x \in A \wedge x \in B \vee x \in A \wedge x \in C$
6) Therefore, $x \in A \cap B \vee x \in A \wedge x \in C$
7) Therefore, $x \in A \cap B \vee x \in A \cap C$
8) 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 $(A \cap B) \cup (A \cap C) \subseteq A \cap B \cup C$]
9) We assume $y \in (A \cap B) \cup (A \cap C)$ and show $y \in A \cap B \cup C$
10) Therefore, $y \in A \cap B \vee y \in A \cap C$
11) Therefore, $(y \in A \wedge y \in B) \vee y \in A \cap C)$
12) Therefore, $(y \in A \wedge y \in B) \vee (y \in A \wedge y \in C)$
13) Therefore, $y \in A \wedge (y \in B \vee y \in C)$
14) Therefore, $y \in A \wedge y \in B \cup C$
15) This finishes the proof. Q.E.D.

Ruleset:
$\_ \Rightarrow$ "appropriate"

Output for the Running Example

1) 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
2) We assume $x \in A \cap B \cup C$ and show $x \in (A \cap B) \cup (A \cap C)$
3) Therefore, $x \in A \land x \in B \cup C$
4) Therefore, $x \in A \land (x \in B \lor x \in C)$
5) Therefore, $x \in A \land x \in B \lor x \in A \land$
6) ~~Therefore, $x \in A \cap B \lor x \in A \land x \in$~~
7) Therefore, $x \in A \cap B \lor x \in A \cap C$
8) We are done with the current part of $x \in (A \cap B) \cup (A \cap C))$. [It remains $(A \cap B) \cup (A \cap C) \subseteq A \cap B \cup C$]
9) We assume $y \in (A \cap B) \cup (A \cap C)$ a
10) Therefore, $y \in A \cap B \lor y \in A \cap C$
11) ~~Therefore, $(y \in A \land y \in B) \lor y \in A$~~
12) ~~Therefore, $(y \in A \land y \in B) \lor (y \in$~~
13) Therefore, $y \in A \land (y \in B \lor y \in C$
14) Therefore, $y \in A \land y \in B \cup C$
15) This finishes the proof. Q.E.D.

Ruleset:
*concepts $\in\{0,1\}$ $\land$ eq-Defn$=0$ $\land$ verb$=$true $\Rightarrow$ too-small
*hypintro$=0$ $\land$ eq-Defn$=0$ $\land$ $\cup$-Defn$=0$ $\land$ verb$=$true $\Rightarrow$ too-small
*concepts $\in\{2,3,4\}$ $\land$ $\cup$-Defn $\in\{1,2,3\}$ $\Rightarrow$ too-big
*hypintro $\in\{1,2,3,4\}$ $\land$ concepts $\in\{2,3,4\}$ $\Rightarrow$ too-big
*unm.c.u.$=0$ $\land$ total $\in\{0,1,2\}$ $\cap$-Defn $\in\{1,2\}$ $\land$ close$=$false $\Rightarrow$ too-small
*eq-Defn$\in\{1,2\}\land$verb$=$false$\Rightarrow$ too-big
*eq-Defn$\in\{1,2\}$ $\land$ verb$=$true $\Rightarrow$ app.
*eq-Defn$=0$ $\land$ verb$=$false $\Rightarrow$ app.
*_ $\Rightarrow$ app.

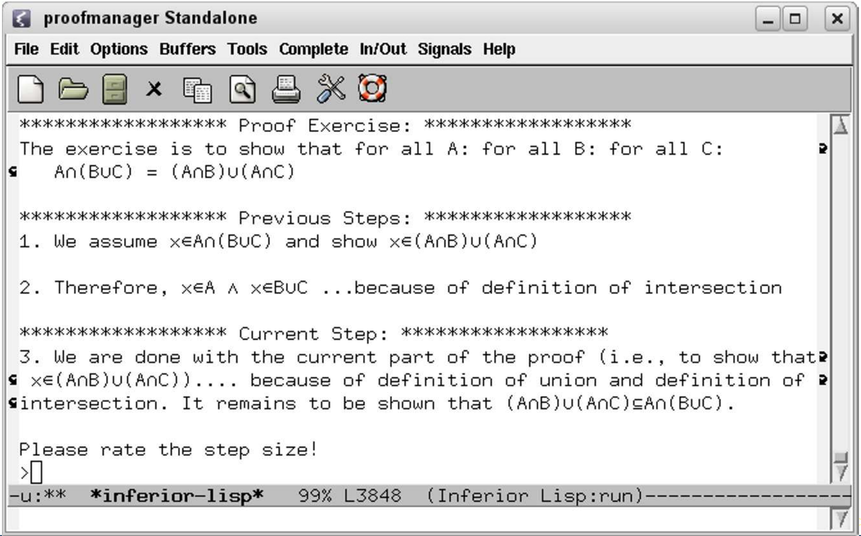
Output for the Running Example

1) 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 $\boxed{7}$,$\boxed{13}$,$\boxed{14}$
2) We assume $x \in A \cap B \cup C$ and show $x \in (A \cap B) \cup (A \cap C)$ $\boxed{1}$
3) Therefore, $x \in A \wedge x \in B \cup C$ $\boxed{2}$
4) Therefore, $x \in A \wedge (x \in B \vee x \in C)$ $\boxed{3}$
5) Therefore, $x \in A \wedge x \in B \vee x \in A \wedge x \in C$ $\boxed{4}$
6) Therefore, $x \in A \cap B \vee x \in A \cap C$ $\boxed{5}$
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 $(A \cap B) \cup (A \cap C) \subseteq A \cap B \cup C$] $\boxed{6}$
8) We assume $y \in (A \cap B) \cup (A \cap C)$ and show $y \in A \cap B \cup C$ $\boxed{8}$
9) Therefore, $y \in A \cap B \vee y \in A \cap C$ $\boxed{9}$
10) Therefore, $y \in A \wedge (y \in B \vee y \in C)$ $\boxed{10}$
11) Therefore, $y \in A \wedge y \in B \cup C$ $\boxed{11}$
12) This finishes the proof. Q.E.D. $\boxed{12}$

Empirical study

- Assess performance of learning classifiers from annotated samples (from different math. domains)
- I.e, proof presentation at different step sizes (1-4 ΩMEGA assertion level steps), (four) human mathematicians judge (label) them as appropriate, too small, too big.
- Apply machine learning (here, J48, PART, SMO, Linear Regression) to learn classifiers, statistical evaluation.

Study Environment



Study Environment

proofmanager Standalone

File Edit Options Buffers Tools Complete In/Out Signals Help

```
****************** Proof Exercise: ******************
The exercise is to show that for all A: for all B: for all C:
   A∩(B∪C) = (A∩B)∪(A∩C)

****************** Previous Steps: ******************
1. We assume x∈A∩(B∪C) and show x∈(A∩B)∪(A∩C)

2. Therefore, x∈A ∧ x∈B∪C ...because of definition of intersection

****************** Current Step: ******************
3. We are done with the current part of the proof (i.e., to show that
 x∈(A∩B)∪(A∩C)).... because of definition of union and definition of
intersection. It remains to be shown that (A∩B)∪(A∩C)⊆A∩(B∪C).

Please rate the step size!
>
-u:**  *inferior-lisp*   99% L3848  (Inferior Lisp:run)----------------
```

step size (w.r.t. assertion level inferences) randomized

explanation randomized

Study Results

- $N_1 = 135$, $N_2 = 198$, $N_3 = 142$, $N_4 = 127$
- most steps rated as appropriate, in particular those generated from one ΩMEGA assertion level step (92%,69%,83%,96%)
- interrater reliability on common sample (61 steps): 71% overall agreement, (Fleiss') multi-rater κ=0.38
- Performance of individually learned classifiers (for each judge), on common sample:

Class. Perf.	Judge 1	Judge 2	Judge 3	Judge 4
% correct	83.6–91.8	68.9–78.7	77.0–90.2	80.3–88.5
Cohen's κ	0.59–0.80	0.44–0.64	0.50–0.78	-0.11–0.23

- Classification more successful for Judge 1 on 135 steps, regression better for Judge 4 ()
- features *total* (# ΩMEGA assertion level steps) and *concepts* alone often predict reasonably well

Conclusion

- explored proof granularity as a classification problem, model/assess different levels of granularity
- assertion level proofs easily yield relevant information for classification and proof presentation task
- proof presentation adapts to samples/different judges, this process is automated.
- future work: application and evaluation in e-learning context