Adaptive Assertion-Level Proofs[1]

Christoph Benzmüller⋆ and Marvin Schiller†

⋆Articulate Software, Angwin, CA, USA
†CeLTech/Saarland University, Germany

EMSQMS
July 20th, 2010

[1] Joint work with Serge Autexier and Dominik Dietrich

Position Statement

'Good' Proofs

- have hierarchical structure (granularity)
- support alternative views
- come together with 'intelligent means'
  - for exploiting the hierarchical structure
  - adapting the presentation of the proof wrt. a given context (user, intention, system resources, ...)

Motivation and Context: Proof Tutoring

Projects OMEGA and DIALOG

OMEGA
(early 90's – recently)

Coarse-grained proofs

- proof plans
- assertion level proofs
- island proofs (sketches)

versus

Proof Grounding/Verification

DIALOG
(early 2000 – recently)

Tutorial Dialog on Proofs

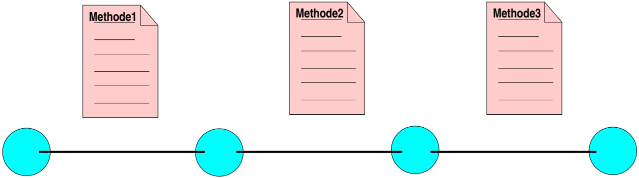
- student proofs
- 'underspecified' steps
- granularity level

versus

Checking of Student Proofs

First papers on assertion level proofs: [HuangEtAl., CADE, 1994], [Huang, CADE, 1994], [HuangFiedler, CADE, 1996], [Meier, CADE 2000]

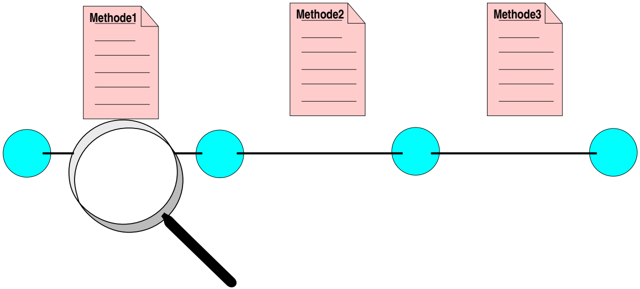
Coarse-grained Proofs



- proof method =

  preconditions-(Tactic)-postconditions

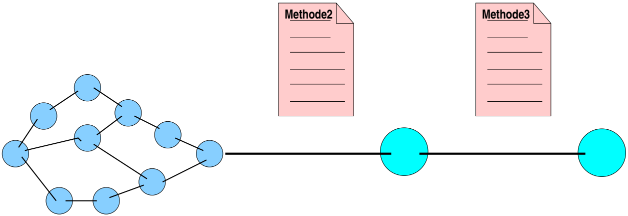
Coarse-grained Proofs



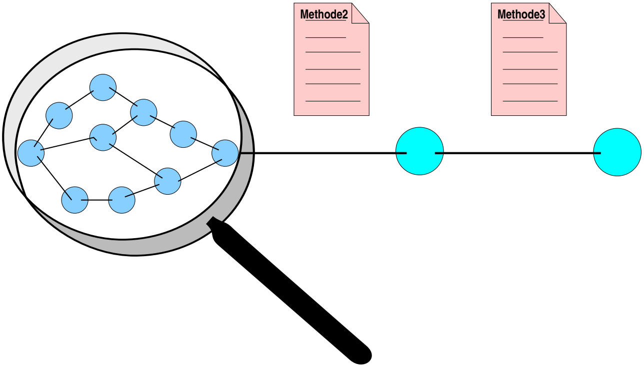
- proof method =

  preconditions-(Tactic)-postconditions
- verification by expansion

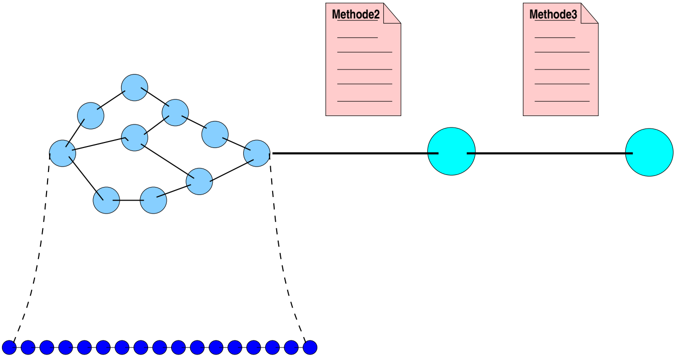
Coarse-grained Proofs



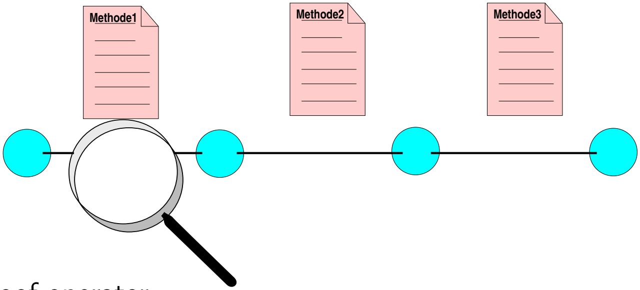
Coarse-grained Proofs



Coarse-grained Proofs



Coarse-grained Proofs

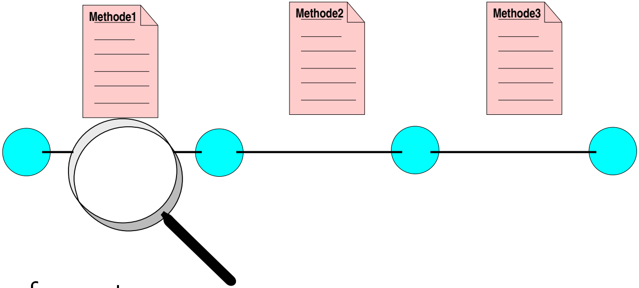


- proof operator =

  preconditions-(Tactic)-postconditions

[Meier, PhD, 2004] [MelisEtAl., AI, 2008]

Coarse-grained Proofs

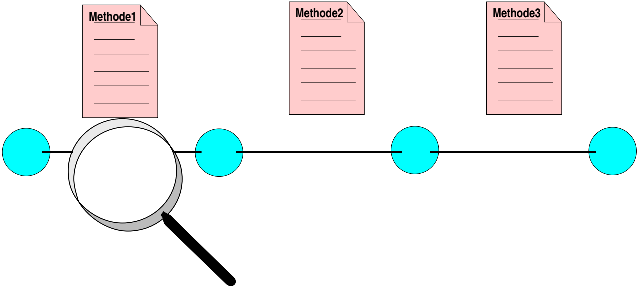


- proof operator =

  preconditions-(ATP)-postconditions

[BenzmuellerEtAl.,J.UCS, 1999]

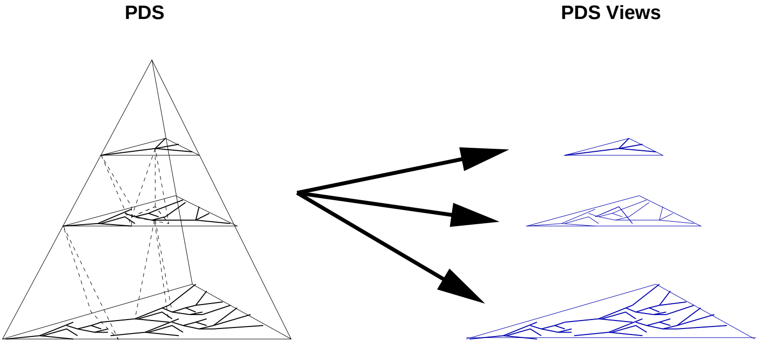
Coarse-grained Proofs



- proof operator =

  preconditions-(AskSomebodyElse)-postconditions

Hierarchical Proof Datastructure (PDS)

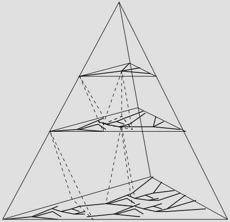


Additionally in PDS of new OMEGA system:

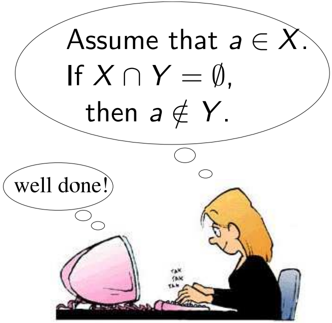
And-Or trees at each layer (proof alternatives)

Projects OMEGA and DIALOG

OMEGA
(early 90's – recently)



DIALOG
(early 2000 – recently)



Overview

1. OMEGA: Assertion Level and Declarative Tactics
2. Tutorial DIALOG and (Adaptive) Proof Granularity
3. Standards for Proof Granularity: Experiments
4. Discussion & Future Work

Assertion Application in ΩMEGA

Inference rules generated from theory

$$\forall A, B, C : \\ A \subseteq B \land B \subseteq C \\ \Rightarrow A \subseteq C$$

Assertion Application in ΩMEGA

Inference rules generated from theory

$$\forall A, B, C: \\ A \subseteq B \wedge B \subseteq C \\ \Rightarrow A \subseteq C \quad \longrightarrow \quad \frac{\mathsf{P}_1 : \mathsf{A} \subseteq \mathsf{B} \quad \mathsf{P}_2 : \mathsf{B} \subseteq \mathsf{C}}{\mathsf{C}_1 : \mathsf{A} \subseteq \mathsf{C}}$$

Assertion Application in ΩMEGA

Inference rules generated from theory

$$\forall A, B, C : A \subseteq B \wedge B \subseteq C \Rightarrow A \subseteq C \quad \longrightarrow \quad \frac{\mathrm{P}_1 : \mathrm{A} \subseteq \mathrm{B} \quad \mathrm{P}_2 : \mathrm{B} \subseteq \mathrm{C}}{\mathrm{C}_1 : \mathrm{A} \subseteq \mathrm{C}}$$

Deep inference application (example)

$$P \Rightarrow (A \subseteq B) \vdash Q \Rightarrow (A \subseteq C)$$

Assertion Application in ΩMEGA

Inference rules generated from theory

$$\forall A, B, C : A \subseteq B \wedge B \subseteq C \Rightarrow A \subseteq C \quad \longrightarrow \quad \frac{\mathrm{P}_1 : \boxed{A \subseteq B} \qquad \mathrm{P}_2 : \mathrm{B} \subseteq \mathrm{C}}{\mathrm{C}_1 : \boxed{A \subseteq C}}$$

Deep inference application (example)

$$P \Rightarrow \boxed{(A \subseteq B)} \vdash Q \Rightarrow \boxed{(A \subseteq C)}$$

With the above mapping, $P_2 \rightarrow B \subseteq C$.

Resulting sequent: $P \Rightarrow (A \subseteq B) \vdash Q \Rightarrow \boxed{(\boxed{P} \wedge (B \subseteq C))}$

Procedural vs. Declarative Proof

- recent trend towards declarative proof languages, inspired by MIZAR

procedural style

```
theorem natcomp:
        "(a::nat) + b =
b+a"
  apply (induct a)
  apply (subst add_0)
  apply (subst add_0_right)
  apply (rule refl)
  apply (subst
add_Suc_right)
  apply (subst add_Suc)
  apply (simp)
done
```

declarative style

```
theorem natcomplus:
        "(a::nat) + b = b+a"
proof (induct a)
  show "0 + b = b + 0"
  proof (-)
    have "0+b=b" by (simp)
    also have "...=b+0" by
(simp)
    finally show ?thesis .
  qed
  next ...
```

[Autexier & Dietrich, ITP 2010]

Basic Declarative Tactics

```
strategy natinduct
cases * ⊢ P x
  with x in (analyzeinductvars "P")
->
proof
  subgoals by (induct x)
    subgoal P 0
    subgoal P (suc x) using IH: P x
  end
```

precondition

action

- make **context** available via precondition
- allow for **internal computations**
- **schematic proof script** as body

Realization

- define tactic language on top of **proof language**
- justification is a **declarative proof script**

[Autexier & Dietrich, ITP 2010]

Granularity

- State-of-art systems generate and maintain various levels of granularity; e.g. proof planners (e.g. λCLAM/HiProofs, Multi/ΩMEGA) etc.
- Granularities result from particular calculi, mechanisms, tactics
- Generally (and in particular, for maths teaching), need to determine appropriate granularity

We consider two applications for granularity judgments

- Automated assessment of student's proofs
- Generation of proof presentions at adapted levels of granularity

The DIALOG Project

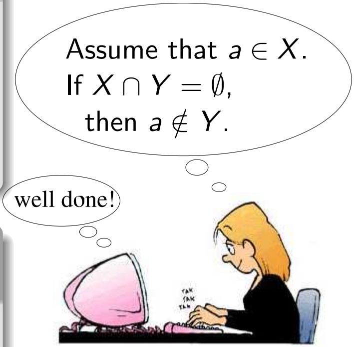
Tutorial Dialog for Mathematics.

Employed Techniques

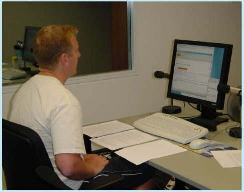
- Dyn. domain reasoning for math proofs: math assistant ΩMEGA
- NL processing, dialogue management, teaching model

Research Processes (in spiral model)

- DIALOG system design
- Prototype development
- Empirical studies



Simulation (Wizard-of-Oz)



Let R and S be relations in a set M. It holds that: $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$. Do the proof interactively with the system!

A pair (x, y) is element of $R \circ S$ iff there is a z in M such that $(x, z) \in R$ and $(z, y) \in S$

Correct!

Therefore a pair (x, y) is element of $(R \circ S)^{-1}$ if there is a z in M, such that $(x, z) \in S$ and $(z, y) \in R$

That's not correct!

Student's Proof

Ex: Show
$(R \circ S)^{-1} = S^{-1} \circ R^{-1}$!

Proof Reconstruction using ΩMEGA

Assertion Level Proof

Student's Proof

Ex: Show
$(R \circ S)^{-1} = S^{-1} \circ R^{-1}$!

s1: $\boxed{\text{Let } (x,y) \in (R \circ S)^{-1}}$.

Proof Reconstruction using ΩMEGA

Assertion Level Proof

Student's Proof

Ex: Show $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$!

s1: Let $(x, y) \in (R \circ S)^{-1}$.

s2: Hence $(y, x) \in (R \circ S)$.

s2: $(y, x) \in (R \circ S) \vdash (x, y) \in \Theta$

Def. $^{-1}$

Christoph Benzmüller* and Marvin Schiller†

Proof Reconstruction using ΩMEGA

Assertion Level Proof

Student's Proof

Ex: Show $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$!

s1: Let $(x, y) \in (R \circ S)^{-1}$.

s2: Hence $(y, x) \in (R \circ S)$.

s3: Hence $(y, z) \in R \wedge (z, x) \in S$.

$$\text{s3}: (y, z) \in R \wedge (z, x) \in S \vdash (x, y) \in \Theta$$

Def.$\circ$

$$\text{s2}: (y, x) \in (R \circ S) \vdash (x, y) \in \Theta$$

Def.$^{-1}$

Christoph Benzmüller* and Marvin Schiller†

Proof Reconstruction using ΩMEGA

Assertion Level Proof

Student's Proof

Ex: Show $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$!

s1: Let $(x, y) \in (R \circ S)^{-1}$.

s2: Hence $(y, x) \in (R \circ S)$.

s3: Hence $(y, z) \in R \wedge (z, x) \in S$.

s4: Hence $(z, y) \in R^{-1} \wedge (x, z) \in S^{-1}$.

$$\vdots$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\text{s4}: (z,y) \in R^{-1} \wedge (x,z) \in S^{-1} \vdash (x,y) \in \Theta}{(y,z) \in R \wedge (x,z) \in S^{-1} \vdash (x,y) \in \Theta}\ \mathbf{Def}^{-1}}{\text{s3}: (y,z) \in R \wedge (z,x) \in S \vdash (x,y) \in \Theta}\ \mathbf{Def}.^{-1}}{\text{s2}: (y,x) \in (R \circ S) \vdash (x,y) \in \Theta}\ \mathbf{Def}.\circ}{\cdots}\ \mathbf{Def}.^{-1}$$

Tutors Analyze Student's Pace

Student	Tutor
	Exercise: $(R \circ S)^{-1} = (x,y) \in S^{-1} \circ R^{-1}$
$\vdots$	
$(x,y) \in (R \circ S)^{-1}$	
	Now try to draw conclusions from this! correct | appropriate | relevant
$(x,y) \in S^{-1} \circ R^{-1}$	
	This cannot be concluded directly. You need some intermediate steps! correct | too coarse-grained | relevant

Granularity: The question of the appropriate proof step size.

Granularity as a Classification Problem

- We consider composite proof steps as **aggregations** of inference steps (which may potentially be unfolded into intermediate steps)
- Assign labels to single-inference or composite proof steps; **appropriate**, **too small**, or **too big**.
- Models for granularity: classifiers (mappings *criteria* $\Rightarrow$ *verdict*)

Analysis of Proof Steps as Basis for Classification

Granularity Criteria

- Content: Which and how many concepts are employed? What (mathematical) theories do they belong to? Are definitions, theorems or lemmata employed?
- Structural properties: New hypotheses or subgoals? Is a step similar to a sequence of previous steps? Are the manipulations restricted to the same formula part? Direction of inference?
- User knowledge: Are the employed concepts known to the user?
- Explicitness: Are the employed concepts named explicitly?

Analysis results are represented as a vector of observations for each step and encoded numerically

Example

Student step	Infs	Features	Verdict
1. We assume $(y,x) \in (R \circ S)^{-1}$ and show $(y,x) \in S^{-1} \circ R^{-1}$	Def.=, Def.$\subseteq$	total:2, concepts:2, relations:0, verb:0, ...	?
2. Hence, $(x,y) \in R \circ S$	Def^{-1}	total:1, concepts:1, relations:1, verb:0, ...	?

...

Example

Student step	Infs	Features	Verdict
1. We assume $(y,x) \in (R \circ S)^{-1}$ and show $(y,x) \in S^{-1} \circ R^{-1}$	Def.=, Def.$\subseteq$	total:2, concepts:2, relations:0, verb:0, ...	appropriate
2. Hence, $(x,y) \in R \circ S$	Def^{-1}	total:1, concepts:1, relations:1, verb:0, ...	appropriate

...

Sample ruleset classifier

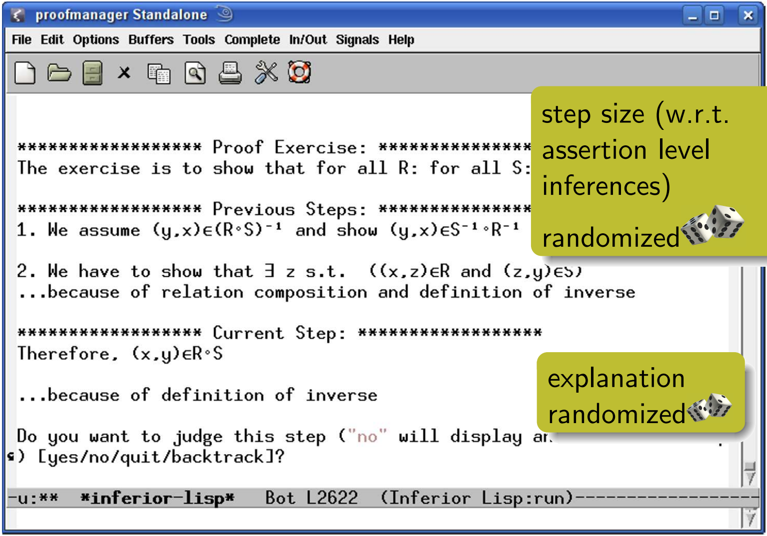
* total $\in \{0,1,2\} \Rightarrow$ appropriate
* unmastered $\in \{2,3,4\} \wedge$ relations $\in \{2,3,4\} \Rightarrow$ too big
* total $\in \{3,4\} \wedge$ relations $\in \{0,1\} \Rightarrow$ too big
* unmastered $\in \{0,1\} \Rightarrow$ appropriate
* _ $\Rightarrow$ appropriate

Standards for Proof Granularity

Experiments

- Can we learn how to judge granularity from experts? What can we learn?
- Four experts with teaching experience
- Presented proof steps constructed at various sizes (aggregated from 1-3 assertion level inference applications), collected granularity judgments (using experiment environment)
- Analyzed (raw) data and learned classifiers (via PART, J48, SMO)

Experiment Environment



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

randomized

explanation randomized

Experiment Results

- Experts agree "moderately" (multirater variation of Brennan and Prediger's $\kappa = 0.57$)
- When experts agree:
  appropriate steps: **one** or **two** ass. level inf. appl.
  too big steps: **three** ass. level inf. appl.
- Large majority of presented steps judged **appropriate**
- Judgments by individual tutors: more elaborate classifiers learned (but results differ among experts)

Discussion

- Tutoring of mathematical proofs: not only correctness, but also granularity (and relevance) play a role
- Approach requires proofs with “meaningful” information (structure, concept ontology), no “black box tactics”

Future Work

- Investigate differences in granularity across further
  - application domains of automated proofs
  - mathematical domains
  - communities
  - levels of expertise
- Compare usefulness of input attributes delivered by OMEGA and other systems

Proof Presentation from Assertion Level Proof

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"

Proof Presentation from Assertion Level Proof

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,~~ $\cancel{x \in A \cap B \vee x \in A \wedge x \in C}$
7) Therefore, $x \in A \cap B \vee x \in A \cap C$
8) We are done with the current part of the
$B) \cup (A \cap C))$. [It remains to be shown tl
9) We assume $y \in (A \cap B) \cup (A \cap C)$ and
10) Therefore, $y \in A \cap B \vee y \in A \cap C$
11) ~~Therefore,~~ $\cancel{(y \in A \wedge y \in B) \vee y \in A \cap}$
12) ~~Therefore,~~ $\cancel{(y \in A \wedge y \in B) \vee (y \in A}$
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:

* Hypintro=1 $\wedge$ total$>$ 1 $\Rightarrow$ step-too-big
* $\cup$-Defn $\in 1, 2 \wedge \cap$-Defn $\in 1, 2$ $\Rightarrow$ step-too-big
* $\cap$-Defn$<$ 3 $\wedge$ $\cup$-Defn=0 $\wedge$ masteredconceptsunique=1 $\wedge$ unmasteredconceptsunique=0 $\Rightarrow$ step-too-small
* _ $\Rightarrow$ step-appropriate