Generation of hints, next steps and complete solutions for axiomatic Hilbert style proofs

Josje Lodder
31-05-2016

Open Universiteit
www.ou.nl

Research questions

- How can we provide feedback and feedforward in e-learning tools to support students with their tasks in logic
- How effective are these tools?

- We restrict these questions to the following subjects
  - standard equivalences and normal forms
  - Hilbert style axiomatic proofs
  - structural induction

topic of this talk

Open Universiteit
www.ou.nl

Hilbert style axiomatic proofs

To prove $\Sigma \vdash \varphi$ you can use :

- 3 axioms:

  A $\varphi \to (\psi \to \varphi)$

  B $(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$

  C $(\neg \psi \to \neg \varphi) \to (\varphi \to \psi))$

- Assumptions

  $\varphi \vdash \varphi$

- Modus Ponens

  If $\Phi \vdash \varphi$ and $\Delta \vdash \varphi \to \psi$ then $\Phi, \Delta \vdash \psi$

- Deduction theorem

  If $\Sigma, \varphi \vdash_S \psi$ then $\Sigma \vdash_S \varphi \to \psi$

Open Universiteit
www.ou.nl

Example proof

Proof for $p \to (q \to r), p \to q \vdash_S p \to r$

1	$p \to (q \to r) \vdash_S p \to (q \to r)$	assumption
2	$\vdash_S (p \to (q \to r)) \to ((p \to q) \to (p \to r))$	axiom b
3	$p \to (q \to r) \vdash_S (p \to q) \to (p \to r)$	MP 1, 2
4	$p \to q \vdash_S p \to q$	assumption
5	$p \to (q \to r), p \to q \vdash_S p \to r$	MP 3, 4

Open Universiteit
www.ou.nl

Alternative proof

Proof of $p \to (q \to r), p \to q \vdash_S p \to r$

1.	$p \to (q \to r) \vdash_S p \to (q \to r)$	assumption
2.	$p \vdash_S p$	assumption
3.	$p \to (q \to r), p \vdash_S q \to r$	MP 1, 2
4.	$p \to q \vdash_S p \to q$	assumption
5.	$p \to q , p \vdash_S q$	MP 2, 4
6.	$p \to (q \to r), p \to q , p \vdash_S r$	MP 3, 5
7.	$p \to (q \to r), p \to q \vdash_S p \to r$	Deduction 6

Open Universiteit
www.ou.nl

Desired features of an e-learning tool for stepwise exercises

- Stepwise solution of an exercise
- Feedback on mistakes
  - syntactical mistakes
  - rule mistakes
  - strategic mistakes
- Hints and next steps
- Complete solutions

you need a solution strategy !

Open Universiteit
www.ou.nl

Strategies for constructing axiomatic proofs

- Constructive completeness proof:
  - produces extremely long proofs
- Translation of semantic tableau method in axiomatic proof (Harrison)
  - only indirect proofs: to prove $\Sigma \vdash \varphi$, show first: $\Sigma, \neg\varphi \vdash \bot$

Natural deduction

$$\frac{\begin{array}{l}\varphi \ (1) \\ \ldots \\ \psi\end{array}}{\varphi \rightarrow \psi \ (-1)}$$

Deduction theorem

$$\frac{\varphi \rightarrow \psi \qquad \varphi}{\psi}$$

Modus ponens

$$\frac{\neg\neg\varphi}{\varphi}$$

$$\frac{\begin{array}{ll}\psi \ (1) & \psi \ (1) \\ \ldots & \ldots \\ \varphi & \neg\varphi\end{array}}{\neg\psi \ (-1)}$$

Strategy for constructing natural deduction proofs (Bolotov)

Find a proof of $\Sigma \vdash \varphi$ via a set of transformations of $\Sigma' \vdash \Delta, \varphi$

where Σ' is the current set of assumptions, and Δ, φ a stack of goals.

Transformations:

- $\Sigma \vdash \Delta, p \quad \Rightarrow \quad \Sigma, \neg p \vdash \Delta, p,$ false
- $\Sigma \vdash \Delta, \neg\varphi \quad \Rightarrow \quad \Sigma, \varphi \vdash \Delta, \neg\varphi,$ false
- $\Sigma \vdash \Delta, \varphi \rightarrow \psi \quad \Rightarrow \quad \Sigma, \varphi \vdash \Delta, \varphi \rightarrow \psi, \psi$

Before adding a new goal, check whether the current goal is reached, by applying modus ponens and double negation to the set of assumptions and reached goals.

Strategy for constructing natural deduction proofs (Bolotov) (2)

- If no rules are applicable use assumptions:

  - $\Sigma, \neg\varphi \vdash \Delta, \text{false} \quad \Rightarrow \quad \Sigma, \neg\varphi \vdash \Delta, \text{false}, \varphi$
  - $\Sigma, \varphi \to \psi \vdash \Delta, \text{false} \quad \Rightarrow \quad \Sigma, \varphi \to \psi \vdash \Delta, \text{false}, \varphi$

Strategy for axiomatic proofs

We use

- a stack of goals: $\Sigma \vdash \varphi$
- a set of availables *A*: prooflines: nr, $\Sigma \vdash \varphi$ (motivation, [nrs])
- a partial proof *P*

Repeat the following steps:

- Close *A* under modus ponens and double negation
- Check if a goal is reachable: delete reached goals from the stack and add them to *A*
- Add the subproof of this goal to *P.*
- Use the transformation rules to create new goals, add the new goal to *P*
- Use the assumptions to create new goals if no other rule is applicable

- Use predefined subproofs for double negation and contradiction rules.

Open Universiteit
www.ou.nl

Example

goals:

p -> q, p -> (q -> r) |- p -> r

availables:

1. *p -> q |- p -> q* — *Assumption*
2. *p -> (q -> r) |- p -> (q -> r)* — *Assumption*

proof:

1. *p -> q, p -> (q -> r) |- p -> r* — *?*

=> deduction

start

goals:

p, p -> q, p -> (q -> r) |- r , p -> q, p -> (q -> r) |- p -> r

availables:

1. *p -> q |- p -> q* — *Assumption*
2. *p -> (q -> r) |- p -> (q -> r)* — *Assumption*
3. *p |- p* — *Assumption*

proof:

2. *p, p -> q, p -> (q -> r) |- r* — *?*
1. *p -> q, p -> (q -> r) |- p -> r* — *Deduction, 2*

result after an application of deduction

Example (continued)

=> modus-ponens

goals:

p, p -> q, p -> (q -> r) |- r

p -> q, p -> (q -> r) |- p -> r

availables:

1.	p -> q |- p -> q	Assumption
2.	p -> (q -> r) |- p -> (q -> r)	Assumption
3.	p |- p	Assumption
4.	p, p -> q |- q	Modus Ponens, 1, 3
5.	p, p -> (q -> r) |- q -> r	Modus Ponens, 2, 3
6.	p, p -> q, p -> (q -> r) |- r	Modus Ponens, 5, 4

Goal is reached

proof:

2.	p, p -> q, p -> (q -> r) |- r	?
1.	p -> q, p -> (q -> r) |- p -> r	Deduction, 2

Example (completion)

=> impl-intro

goals:

availables:

1. p -> q |- p -> q   Assumption
2. p -> (q -> r) |- p -> (q -> r)   Assumption
3. p |- p   Assumption
4. p, p -> q |- q   Modus Ponens, 1, 3
5. p, p -> (q -> r) |- q -> r   Modus Ponens, 2, 3
6. p, p -> q, p -> (q -> r) |- r   Modus Ponens, 5, 4
7. p -> q, p -> (q -> r) |- p -> r   Deduction, 6

proof:

3. p -> q |- p -> q   Assumption
4. p -> (q -> r) |- p -> (q -> r)   Assumption
5. p |- p   Assumption
6. p, p -> q |- q   Modus Ponens, 3, 5
7. p, p -> (q -> r) |- q -> r   Modus Ponens, 4, 5
2. p, p -> q, p -> (q -> r) |- r   Modus Ponens, 7, 6
1. p -> q, p -> (q -> r) |- p -> r   Deduction, 2

Add heuristics

Now we can produce proofs, but these proofs use the axioms only in subproofs concerning negations or contradicions. An e-learning tool should also help students to recognize applicable axioms.

Therefore we introduce heuristics:

In the step;

Close *A* under modus ponens and double negation

add:

applicable/useful versions of axiom A, axiom B and axiom C

Example: if goal $= \Sigma \vdash \varphi \rightarrow \psi$ and $\Sigma \vdash \neg\varphi$ in availables,

add instances to the availables::

$\vdash \quad \neg\varphi \rightarrow (\neg\psi \rightarrow \neg\varphi)$ (axiom A)

$\vdash \quad (\neg\psi \rightarrow \neg\varphi) \rightarrow (\varphi \rightarrow \psi)$ (axiom C)

Open Universiteit
www.ou.nl

Example

availables:

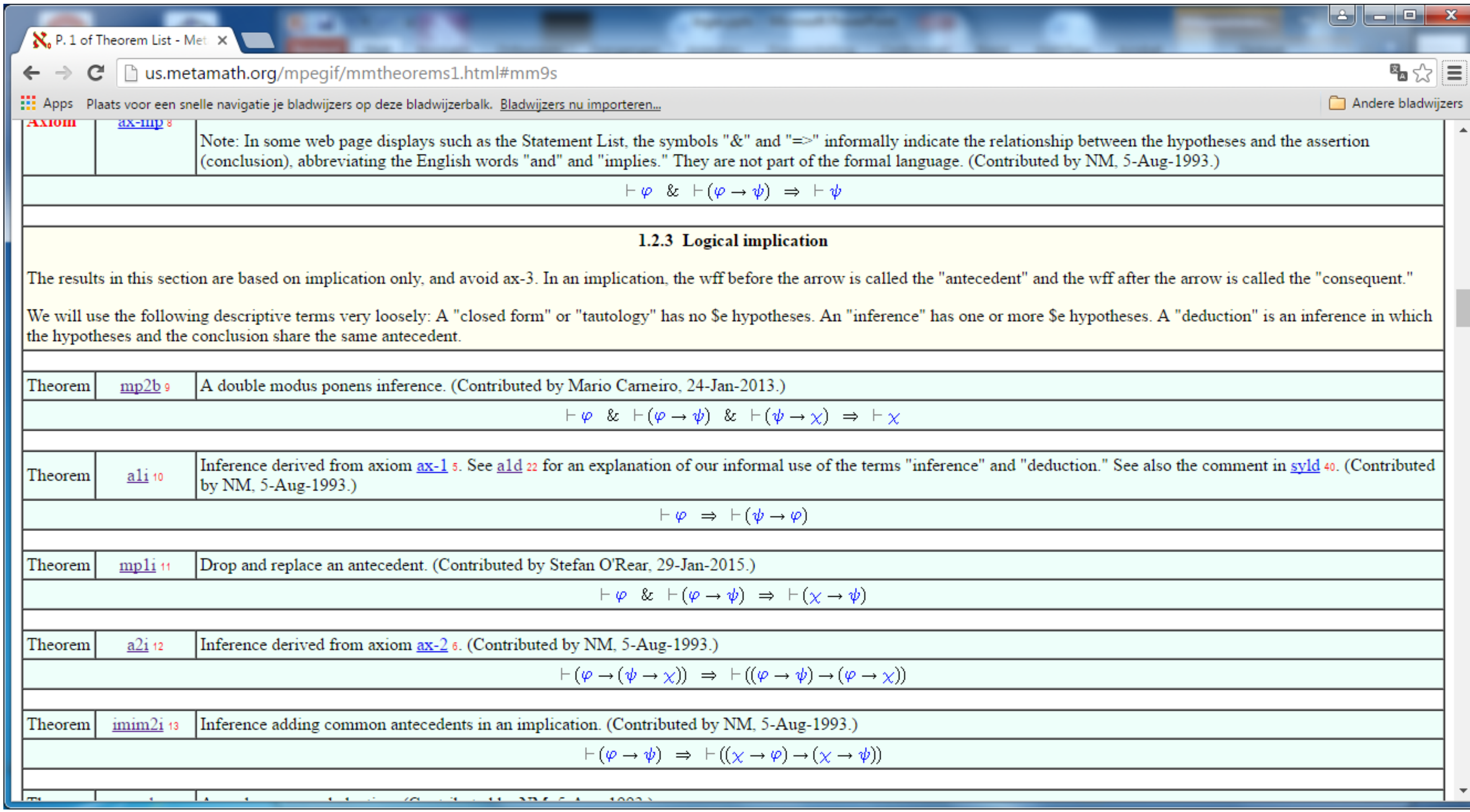
1. p -> q |- p -> q	Assumption
2. p -> (q -> r) |- p -> (q -> r)	Assumption
3. |- (p -> (q -> r)) -> ((p -> q) -> (p -> r))	Axiom b
4. p -> (q -> r) |- (p -> q) -> (p -> r)	Modus Ponens, 3, 2
5. p -> q, p -> (q -> r) |- p -> r	Modus Ponens, 4, 1

proof:

2. p -> q |- p -> q	Assumption
3. p -> (q -> r) |- p -> (q -> r)	Assumption
4. |- (p -> (q -> r)) -> ((p -> q) -> (p -> r))	Axiom b
5. p -> (q -> r) |- (p -> q) -> (p -> r)	Modus Ponens, 4, 3
1. p -> q, p -> (q -> r) |- p -> r	Modus Ponens, 5, 2

How good is the strategy (1)?

- Comparison with metamath proof list:



P. 1 of Theorem List - Met

us.metamath.org/mpegif/mmtheorems1.html#mm9s

Apps Plaats voor een snelle navigatie je bladwijzers op deze bladwijzerbalk. Bladwijzers nu importeren... Andere bladwijzers

Axiom	ax-mp 8	Note: In some web page displays such as the Statement List, the symbols "&" and "=>" informally indicate the relationship between the hypotheses and the assertion (conclusion), abbreviating the English words "and" and "implies." They are not part of the formal language. (Contributed by NM, 5-Aug-1993.)
		$\vdash \varphi \quad \& \quad \vdash (\varphi \to \psi) \quad \Rightarrow \quad \vdash \psi$

1.2.3 Logical implication

The results in this section are based on implication only, and avoid ax-3. In an implication, the wff before the arrow is called the "antecedent" and the wff after the arrow is called the "consequent."

We will use the following descriptive terms very loosely: A "closed form" or "tautology" has no $e hypotheses. An "inference" has one or more $e hypotheses. A "deduction" is an inference in which the hypotheses and the conclusion share the same antecedent.

Theorem	mp2b 9	A double modus ponens inference. (Contributed by Mario Carneiro, 24-Jan-2013.)
		$\vdash \varphi \quad \& \quad \vdash (\varphi \to \psi) \quad \& \quad \vdash (\psi \to \chi) \quad \Rightarrow \quad \vdash \chi$
Theorem	a1i 10	Inference derived from axiom ax-1 5. See a1d 22 for an explanation of our informal use of the terms "inference" and "deduction." See also the comment in syld 40. (Contributed by NM, 5-Aug-1993.)
		$\vdash \varphi \quad \Rightarrow \quad \vdash (\psi \to \varphi)$
Theorem	mp1i 11	Drop and replace an antecedent. (Contributed by Stefan O'Rear, 29-Jan-2015.)
		$\vdash \varphi \quad \& \quad \vdash (\varphi \to \psi) \quad \Rightarrow \quad \vdash (\chi \to \psi)$
Theorem	a2i 12	Inference derived from axiom ax-2 6. (Contributed by NM, 5-Aug-1993.)
		$\vdash (\varphi \to (\psi \to \chi)) \quad \Rightarrow \quad \vdash ((\varphi \to \psi) \to (\varphi \to \chi))$
Theorem	imim2i 13	Inference adding common antecedents in an implication. (Contributed by NM, 5-Aug-1993.)
		$\vdash (\varphi \to \psi) \quad \Rightarrow \quad \vdash ((\chi \to \varphi) \to (\chi \to \psi))$

Proofs without deduction theorem

- Use the proof of the deduction theorem to rewrite proofs with deduction in proofs without this rule.
- Apply this rewriting only in necessary cases
- Clean up rewritten proofs.

- Simple rewriting the first example proof (with deduction) produces a 20 line proof, ‘smart’ rewriting produces our second 5-line proof.

Open Universiteit
www.ou.nl

Comparison metamath-org

thm	#metamath	#deduction	#smartnodeduction
mp2b	5	5	5
ali	3	2	3
mpli	5	4	5
a2i	3	3	3
imim2i	5	7	5
mpd	5	7	5
syl	7	6	7
mpi	7	6	7
id1	5	2	5
a1d	7	5	7
a2d	7	6	7
sylcom	9	10	9
syl5com	15	9	15
com12	9	8	9
syl5	23	9	19
syl6	11	9	11
pm2.27	13	5	13
mpdd	11	9	11
mpid	17	11	17
pm2.43i	9	5	9
pm2.43a	11	9	11
pm2.43	15	6	11
imim2d	13	10	13
imim2	7	8	7

results until now:

- 24 proofs compared
- 22 proofs up to order equal to our proofs
- 2 shorter proofs

Open Universiteit
www.ou.nl

How good is the strategy (2)

- Compare the generated proof with student solutions

- Can we use this strategy to provide hints/next steps

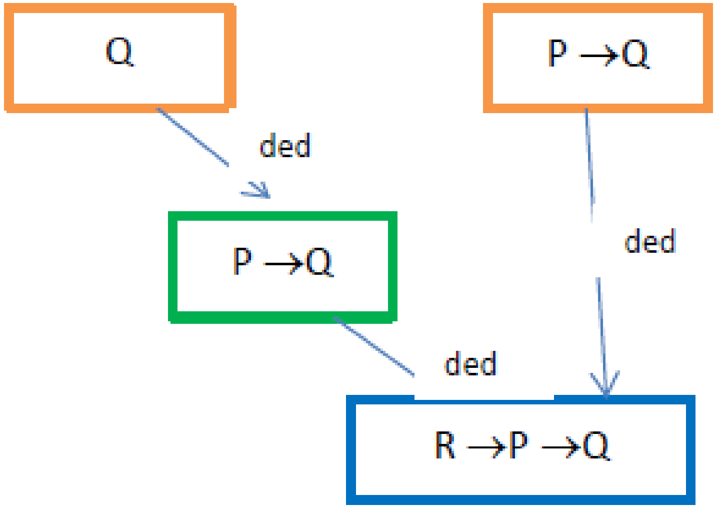
Open Universiteit
www.ou.nl

Linear proofs vs proof DAGs

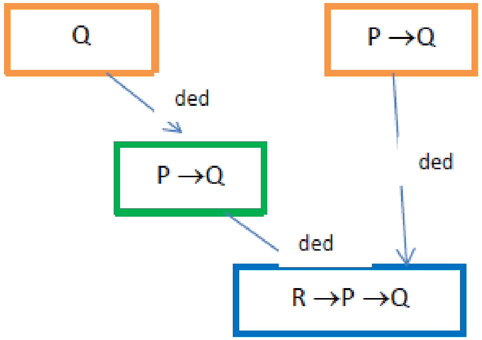
- We can use this strategy to generate complete proofs, and we could also use it to give hints and next steps, also if a student constructs a different solution, by adding the steps of the student to the availables
- We cannot use it within the IDEAS frame work to monitor the steps of the student.
- Our solution:
  - the availables form a proof-DAG from which we can extract linear proofs
  - we expand the availables
  - from this extended proof-DAG we construct a non-deterministic strategy which produces different solutions
  - we can use this strategy to recognize the steps of the student

Open Universiteit
www.ou.nl

Example DAG



Example of strategygeneration



Derivation #8

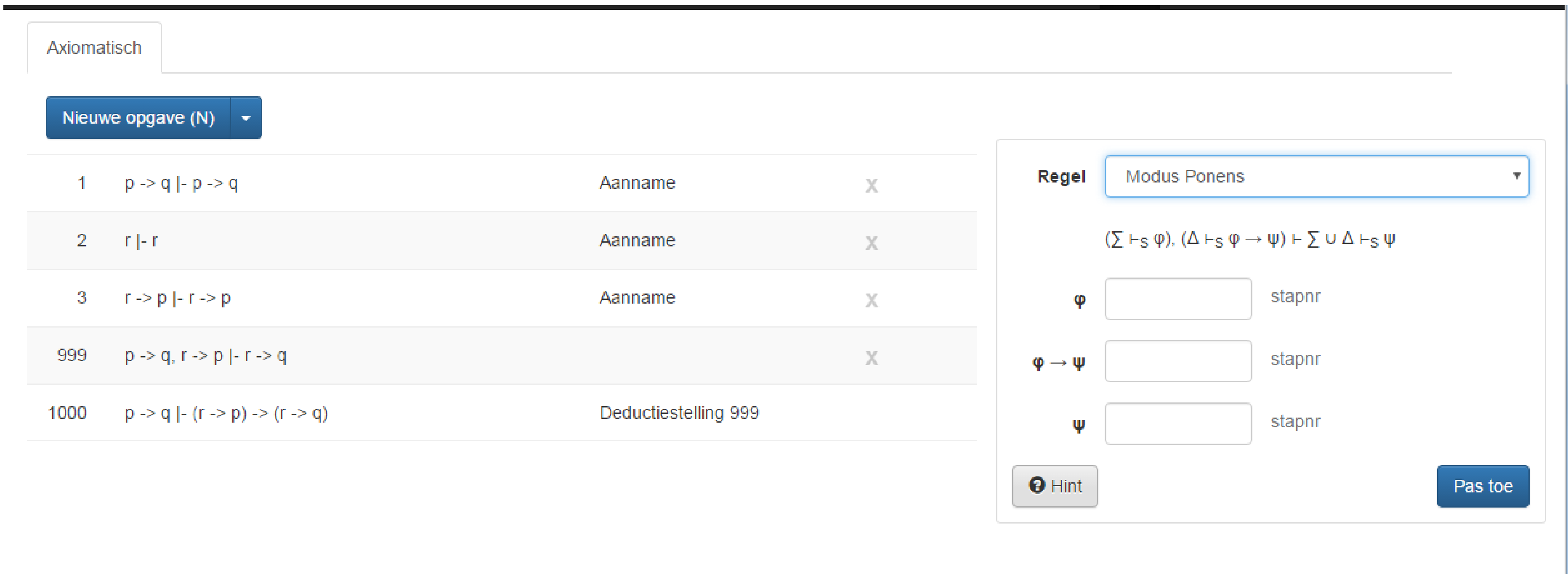
=> as
3. "P->Q" As

=> ded
3. "P->Q" As
4. "R->P->Q" Ded, 3

Derivation #7

=> as
1. "Q" As
=> ded
1. "Q" As
2. "P->Q" Ded, 1
=> ded
1. "Q" As
2. "P->Q" Ded, 1
4. "R->P->Q" Ded, 2

LogAx



Open Universiteit
www.ou.nl

Hints, next steps and feedback in LogAx

- Use the strategy to provide different level hints:
  - goal
  - rule
  - next step

- Use the collection of common mistakes to provide feedback

- Evaluation:
  - do students learn from using LogAx?
  - do students make less mistakes in choosing applicable rules?
  - do students make more mistakes in correctly applying rules?

Open Universiteit
www.ou.nl

Dank voor jullie aandacht!

Open Universiteit
www.ou.nl