Using a proof assistant in an introduction to proof course: first experiment with two proof assistants

Evmorfia-Iro Bartzia, Pierre Boutry, and Julien Narboux

2 July 2024

- Project APPAM (French Agency of Research)
- Pre-experiment for a 1st year undergraduate course
- 2 proof assistants (PAs)
- Observation and survey on students perceptions

- Context
- Observations
- Survey
- Questions for further work

The course

- **Title** : Introduction to reasoning and proof
- **Level** : First year undergraduate students
- **Objective**: math vocabulary, reasoning rules, write proofs
- **Content**: naive set theory, functions, relations, natural numbers, induction
- **Organisation** : 24h lectures (and exercises) + 10h using PAs
- **Evaluation** : 2 on paper + 1 on PA

The students

- 200 students of mathematics and computer science
- 6 groups of 30-40 students (2 math + 4 CS)
- Courses taken before:
  - Introduction to programming
  - Introduction to mathematics (algebra and analysis)

The teachers

- Types: lecturers in mathematics and CS, PhD candidates in mathematics, secondary school teachers
- Not experts in PAs
- Lab session conducted by us

Lab Sessions

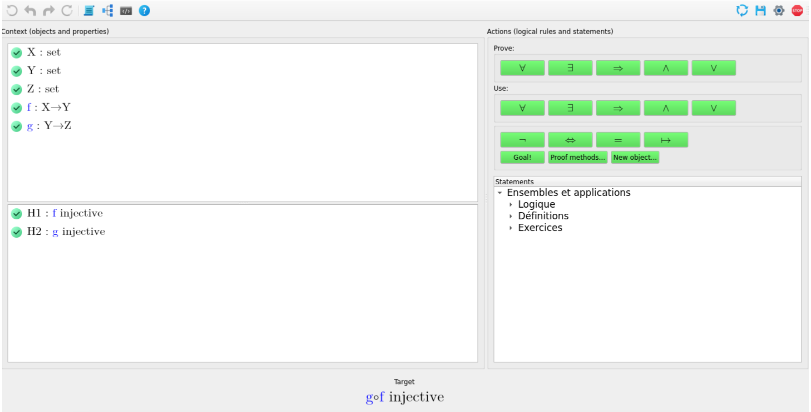
1. **Introduction to the PA** and to propositional logic using exercises about sets
2. **True or False** exercises about basic arithmetic
3. **Functions**: injectivity, surjectivity, direct and inverse image
4. **Functions**: injectivity, surjectivity, direct and inverse image
5. On computer exam.

Verbose L98N

- Alternative language for L98N 4
- developed by Patrick Massot
- *help* feature
- global visualization of the proof

D98duction

- Interface on top of L98N 3
- developed by Frédéric Le Roux
- point-and-click user interface
- local visualization of the proof


Context (objects and properties)
X : set
Y : set
Z : set
f : X→Y
g : Y→Z
H1 : f injective
H2 : g injective
Actions (logical rules and statements)
Prove:
∀ ∃ ⇒ ∧ ∨
Use:
∀ ∃ ⇒ ∧ ∨
¬ ⇔ = ↦
Goal!
Proof methods...
New object...
Statements
Ensembles et applications
Logique
Définitions
Exercices
Target
g∘f injective

Instrumental approach (Rabardel 1995)

Instrumental genesis: the process of an artifact becoming an instrument

- Instrumentalization
- Instrumentation

Instrumentalization

The process of adjusting a specific set of characteristics of a PA to direct students' activity towards a specific aspect of proving.

- using the PA libraries or working in a microworld
- using true or false statements (think semantically)

Instrumentation

The way specific practices of a PA shape the students' process of proving.

- development of automatism about quanti ers
- trial and error strategies

General questions

- What are the possible effects of using PAs on students' learning of proof?
- What characteristics of PAs are likely to strengthen or obstruct these effects?

Question for our work

- How do students perceive their use of PAs?
- Does using one of the two PAs help students ameliorate skills in proof and proving? Which skills?
- Do we observe big differences between the groups? What kind of differences?

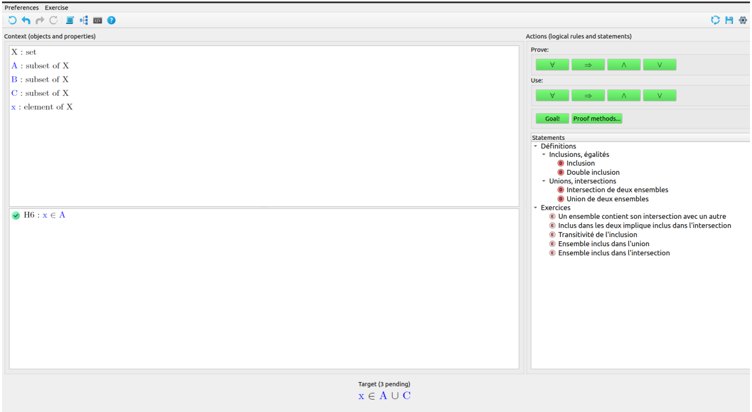
For every a 0, (if for every ϵ 0, a ϵ) then $a = 0$

- **choose the negation**: $: (8a \quad 0, (8\epsilon \quad 0, a \quad \epsilon) \;)\; a = 0)$
- **push negation**: $9a \quad 0, (8\epsilon \quad 0, a \quad \epsilon)$ ^ a 6= 0
- **random value for** a: $8\epsilon \quad 0, 1 \quad \epsilon$
- **decide if it is provable or not**

Exploring conjectures

- Many proceed by trial and error
- They used the PA to generate an example of the statement instantiating with an acceptable value
- Same observation for both PAs

- D*98*duction: local visualization of the proof
- Verbose L*98*N: global visualization of the proof


Preferences Exercise
Context (objects and properties)
X : set
A : subset of X
B : subset of X
C : subset of X
x : element of X
H6 : x ∈ A
Actions (logical rules and statements)
Prove:
∀ ⇒ ∧ ∨
Use:
∀ ⇒ ∧ ∨
Goal!
Proof methods...
Statements
Définitions
Inclusions, égalités
Inclusion
Double inclusion
Unions, intersections
Intersection de deux ensembles
Union de deux ensembles
Exercices
Un ensemble contient son intersection avec un autre
Inclus dans les deux implique inclus dans l'intersection
Transitivité de l'inclusion
Ensemble inclus dans l'union
Ensemble inclus dans l'intersection
Target (3 pending)
x ∈ A ∪ C

D98duction

- The interface acts as blinders
- The effect (positive or negative) depends on the proof
- For students the switch between the local and global view of the proof may be difficult

- Students' perceptions on the use of the PAs
- The survey was optional
- 38 Answers
- Work in progress

Drag and drop questions

Using the PA allowed me to

- make trial and error
- understand what is to be proven
- verify my proof
- to have an overall view of a proof
- to know how to advance in a proof
- write a proof on paper

Open questions

- What you found easy (/di cult) to do in the PA
- What you liked (/didn't like) about the PA
- What was the reason when I was blocked in the PA?

Likert scale questions (1 to 5)

- I have an idea of the proof before starting in the PA.
- I've done proofs in the PA without understanding them.
- I anticipate the result I will get before I click/write a line.
- When I had an idea for a proof, I couldn't do it in the PA.
- I need a paper sheet when working with a PA.
- I think I want to use this PA frequently.
- I felt very con dent using the PA.

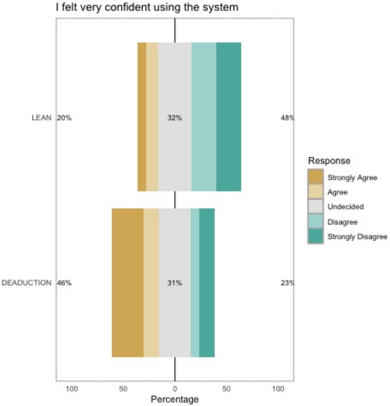
Questions

- What are the students learning?
  - how to use the PA?
  - how to produce proofs using a PA?
  - how to produce proofs in general?
- What is the impact on understanding the concept of proof?
- What is the impact of a PA on writing proofs?

Should take into account different **public**, different **PAs**, different **usage scenarios**, and different **mathematical contents**.

- Analyze the survey results
- Analyze exam results (qualitative analysis)
- Instrument the PA to collect data
- Repeat the experiment

Thank you for your attention!


I felt very confident using the system
LEAN
20%
32%
48%
DEADUCTION
46%
31%
23%
100
50
0
50
100
Percentage
Response
Strongly Agree
Agree
Undecided
Disagree
Strongly Disagree

- Facultative survey – not many answered
- We used the PA as an exerciser – did not work on the translation to paper proofs
- We do not know if they were present during all the lab sessions
- In the meantime they had math courses
- Teachers' different backgrounds (on math, proof, logic, programming) and use of the PA
- Students with different background
- Groups are not homogeneous
- Different proofs on paper and with the PA
- Likert scale

- The assumption hyp is an equality
  One can use it to replace the left-hand-side (namely g (f x1)) by the righ
  (namely g (f x2)) in the goal with:
  `We rewrite using hyp`
- One can use it to replace the right-hand-side in the goal with:
  `We rewrite using ← hyp`
- One can also perform such replacements in an assumption hyp_1 with
  `We rewrite using hyp dans† hyp_1`
- or
  `We rewrite using ← hyp dans† hyp_1`
- One can also use it in a computation step, or combine it linearly to other
  `We combine [hyp, ?_]`
  replacing the question mark by one or more terms proving equalities.

2. $\forall x \in \mathbb{R}\ \exists y \in \mathbb{R}\ x+y>0$
Faux. Montrons que $y=5$ convient pas
On a $x+5>0$. Si $x=-10$. On a $-5>0$. Ce qui est faux.

3. $\exists x \in \mathbb{R}\ \forall y \in \mathbb{R}\ x \leq y^2$
Vrai. Montrons que $x=0$ convient.
On a $0 \leq y^2$ et on sait que $\forall y \in \mathbb{R}$, on a $y^2 \geq 0$. Ce qui prouve la démonstration.

4. $\forall x \in \mathbb{R}\ x^2 \geq x$.
Vrai. Montrons que $x=5$ convient.
On a $5^2 \geq 5$ et $25 \geq 5$. Ce qui est vrai et prouve la démonstration.

4. Supposons que $\forall x \in \mathbb{R}\ x^2 \geq x$

Montrons que cette assertation est vraie avec $x = 0,5$

$x^2 = 0,25$ or $0,25 < 0,5$ $\lightning$

Donc cette assetation est fausse.

3. Supposons que $\exists x \in \mathbb{R}\ \forall y \in \mathbb{R}\ x \leq y^2$

Montrons que $x = 0$ convient

$\forall y \in \mathbb{R},\ y^2 \geq 0$

x convient, l'assertation est vraie.