Structured derivations in mathematics teaching

Ralph-Johan Back

Dept. of Information Technologies,
Abo Akademi University

Cadgme 2009 Linz

Proofs in high school math today

- Proofs are of central importance for understanding mathematics
- But writing proofs is seen as difficult in (Finnish) high school math today, and is usually avoided
- Students are not taught how to write proofs. The few proofs that are given are intuitive and informal, and varying in style and rigour.
  - The best students will still pick out the basic idea of doing mathematics: to derive a result by a sequence of well-justified steps.
  - The rest will see mathematics as just a collection of templates to memorize, solving problems with these without any deeper understanding.
- We need to change this situation: teach how to build proofs in practice, how to use logic in proofs, how to organizing proofs and how to check that the proofs are correct

Calculational proof style

- *Calculational proofs* is a way of presenting mathematical derivations and proofs developed by Edsger Dijkstra, Wim Feijen, Nettie van Gasteren and others in the 1980s.
- A calculational proof is essentially a chain of relational statements

$$t_1 \sim t_2 \sim t_3 \sim \ldots \sim t_{n-1} \sim t_n$$

but written in a special way:

$$
\begin{array}{ll}
\bullet & t_1 \\
\sim & \{\text{motivation for why } t_1 \sim t_2 \text{ holds}\} \\
 & t_2 \\
\sim & \{\text{motivation for why } t_2 \sim t_3 \text{ holds}\} \\
\vdots & \\
\sim & \{\text{motivation for why } t_{n-1} \sim t_n \text{ holds}\} \\
 & t_n
\end{array}
$$

Calculational proofs emphasize

- justification of each proof step,
- use of logical notation, and
- use of logical inference rules

Structured derivations

- *Structured derivations* is a further development of the calculational style, by me and Joakim von Wright in the middle of the 90s. Originally presented in book
  - Back & von Wright: Refinement Calculus: A Systematic Introduction, Springer Verlag1998
- Extends calculational proofs with a number of new features, in particular
  - sub derivations,
  - inherited assumptions, and
  - observations
- A different mathematical basis: Gentzen like proof systems rather than equational logic
  - A structured derivation is equivalent to a Gentzen like proof in higher order logic.
  - But much more user friendly

Structured derivations provide a unified proof format

- Unifies the three main proof formats used today:
  - Forward proofs (Hilbert-style, classic Euclidean proofs)
  - Backward proofs (Gentzen-style, natural deduction), and
  - Derivations (Dijkstra-style proofs, algebraic manipulations, calculations)
- Can express any proof in one of these formats as a structured derivation
- Can mix these different proof formats within a single structured derivation
  - each sub problem can be handled with the format that is most appropriate for this task
  - can combine forward reasoning, backward reasoning and calculations within a single proof format

The importance of explicit justifications

Example 21. Compute the tangent of $\frac{17\pi}{3}$.

Teachers solution:

$$\tan\frac{17\pi}{3} = \tan\left(\frac{6\cdot 2\pi + 5\pi}{3}\right) = \tan\left(2\cdot 2\pi + \frac{5\pi}{3}\right) = \tan\frac{5\pi}{3} = \tan\left(2\pi - \frac{\pi}{3}\right) = -\tan\frac{\pi}{3} = -\sqrt{3}$$

Teacher explains each step verbally.

Student looking at example at home

Example 21. Compute the tangent of $\dfrac{17\pi}{3}$.

Teachers solution:

$$\tan\frac{17\pi}{3} = \tan\left(\frac{6\cdot 2\pi + 5\pi}{3}\right) = \tan\left(2\cdot 2\pi + \frac{5\pi}{3}\right) = \tan\frac{5\pi}{3} = \tan\left(2\pi - \frac{\pi}{3}\right) = -\tan\frac{\pi}{3} = -\sqrt{3}$$

The verbal explanations are lost

- student could not follow the explanation,
- was thinking about something else,
- was not even attending the class, etc

Student looking at example at home

Example 1. Compute the tangent of $\frac{17\pi}{3}$.

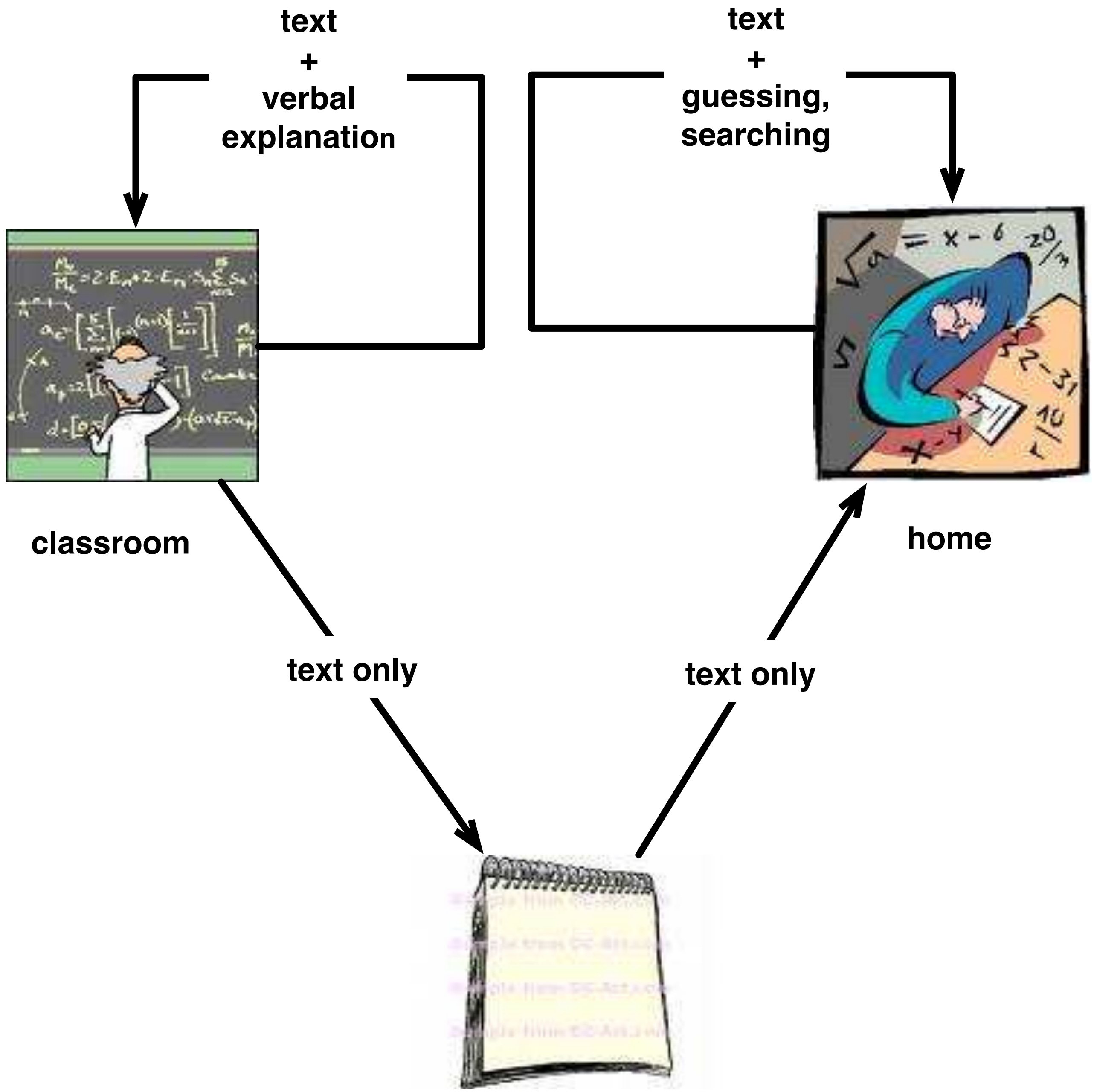
Teachers solution:

$$\tan\frac{17\pi}{3} = \tan\left(\frac{6\cdot 2\pi + 5\pi}{3}\right) = \tan\left(2\cdot 2\pi + \frac{5\pi}{3}\right) = \tan\frac{5\pi}{3} = \tan\left(2\pi - \frac{\pi}{3}\right) = -\tan\frac{\pi}{3} = -\sqrt{3}$$

Students has to decipher this derivation:

- find out motivation for each step
- error prone process, guesswork, trial and error
- wrong guesses can lead to serious problems of understanding

Information loss



Example solution expressed as a structured derivation

- $\tan \frac{17\pi}{3}$

Example solution expressed as a structured derivation

- $\tan \frac{17\pi}{3}$

$=$ {factor out 2π in numerator}

$$\tan\left(\frac{6 \cdot 2\pi + 5\pi}{3}\right)$$

Example solution expressed as a structured derivation

- $\tan\frac{17\pi}{3}$

$=$ $\{$factor out $2\pi\}$

$$\tan\left(\frac{6\cdot 2\pi+5\pi}{3}\right)$$

$=$ $\{$write angle in the form $n\cdot 2\pi+\alpha\}$

$$\tan\left(2\cdot 2\pi+\frac{5\pi}{3}\right)$$

Example solution expressed as a structured derivation

- $\tan \frac{17\pi}{3}$

$=$ {factor out 2π}

$$\tan\left(\frac{6\cdot 2\pi + 5\pi}{3}\right)$$

$=$ {write angle in the form $n\cdot 2\pi + \alpha$}

$$\tan\left(2\cdot 2\pi + \frac{5\pi}{3}\right)$$

$=$ {we can ignore full circles $2\cdot 2\pi$}

$$\tan\frac{5\pi}{3}$$

Example solution expressed as a structured derivation

- $\tan \frac{17\pi}{3}$

$=$ {factor out 2π}

$$\tan\left(\frac{6 \cdot 2\pi + 5\pi}{3}\right)$$

$=$ {write angle in the form $n \cdot 2\pi + \alpha$}

$$\tan\left(2 \cdot 2\pi + \frac{5\pi}{3}\right)$$

$=$ {we can ignore full circles $2 \cdot 2\pi$}

$$\tan \frac{5\pi}{3}$$

$=$ {the angle is in the 4th quadrant, so we can write it in the form $2\pi - \alpha_0$}

$$\tan\left(2\pi - \frac{\pi}{3}\right)$$

$$\tan\left(2\pi - \frac{\pi}{3}\right)$$

$$= \quad \{\text{tangent is negative i 4th quadrant and } \tfrac{\pi}{3} \text{ is basic angle}\}$$

$$-\tan\frac{\pi}{3}$$

$$\tan\left(2\pi - \frac{\pi}{3}\right)$$

$=$ {tangent is negative i 4th quadrant and $\frac{\pi}{3}$ is basic angle}

$$-\tan\frac{\pi}{3}$$

$=$ {tangent of 30 - 60 - 90 triangle}

$$-\sqrt{3}$$

$$\tan\left(2\pi - \frac{\pi}{3}\right)$$

$=$ {tangent is negative i 4th quadrant and $\frac{\pi}{3}$ is basic angle}

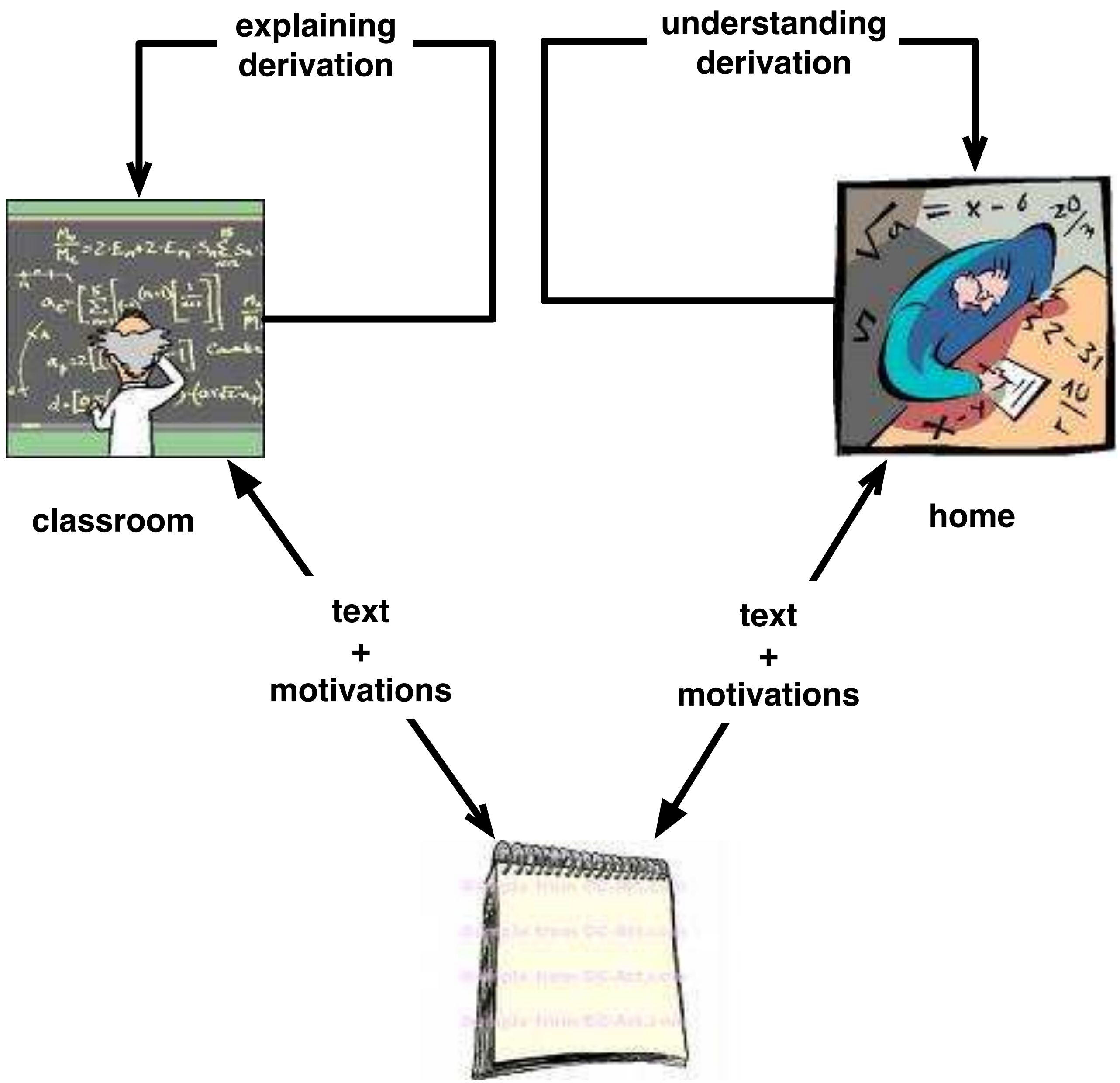
$$-\tan\frac{\pi}{3}$$

$=$ {tangent of 30 - 60 - 90 triangle}

$$-\sqrt{3}$$

□

No information loss



Advantages of this proof format

- Ample space for both terms and for justifications for each step
- Forces the teacher to write the justifications out explicitly, not just stating them verbally
- Forces the student to write out his justifications for each step
- Helps the student to understand the idea of a proof (each step must be justified)
- Student can check his own derivation and find errors in it
- Teacher can identify what the student has not understood
- Easier for teacher to check correctness of students solutions

Structured derivations in teaching

- Structured derivations were originally developed for research purposes (explaining rather complex proofs in programming logic) in 1995 - 2000 (Back and von Wright)
- Turned out to be very useful in teaching mathematics at different educational levels.
- We have been carrying out teaching experiments since 2000 (Back, von Wright, Salakoski, Peltomäki, Mannila, Sibelius, Sallasmaa, Wallin ...):
  - Most teaching experiments carried out in high school math (age group 16 - 18)
  - Experiments on teaching at university freshman level (CS students, age group 18 -)
  - Some experiments on teaching in polytechnic (CS students, age group 19 -)
  - Now starting experiment with teaching structured derivations in junior high (age group 14 - 15)
- Method has been continuously modified and adapted to requirements imposed be teaching in practice

Structured derivations in a nutshell

- **Fixed format for structuring proofs and derivations (a proof language)**
  - teaches the idea of a proof, provides a standard template for how to write down mathematical proofs and derivations in a user friendly way
- **Format for reasoning about the underlying domain is not fixed**
  - can be used for any kind of mathematical reasoning, because the underlying theory does not have to be formalized
- **Use of logical notation and inference rules in proofs**
  - promotes the use of explicit logical notation and logical argumentation in proofs
- **Easy to build computer support for the method**
  - can build editors that support writing structured derivations, with syntax checking, checking correctness of derivations automatically, integration with teaching environments, etc.

Structured derivations syntax

derivation::=

•	task
-	assumption
	⋮
+	justification
	observation
	⋮
⊩	justification
	term
rel	justification
	term
	⋮
□	

justification::=

{motivation}

derivation

derivation

⋮

Logic in high school mathematics

A mathematical proof is a logical argumentation, but logical notation is not used much in high schools, and logical inference rules are seldom given explicitly

When logic is given as a course, it is taught as a separate (mathematical) topic, not as a tool for solving mathematical problems.

What we need is

logical mathematics rather than *mathematical logic*

Logic is abundant in high school math, but not usually recognized

Example 2: Solve the equation

$$(x-1)(x^2+1)=0$$

Example 2

- $$(x-1)(x^2+1)=0$$

Example 2

- $(x-1)(x^2+1) = 0$

$\equiv$ $\{$zero product rule: $ab = 0 \equiv a = 0 \vee b = 0\}$

$x - 1 = 0 \vee x^2 + 1 = 0$

Example 2

• $(x-1)(x^2+1)=0$

$\equiv$ {zero product rule: $ab=0 \equiv a=0 \vee b=0$}

$x-1=0 \vee x^2+1=0$

$\equiv$ {add 1 to both sides in left disjunct}

$x=1 \vee x^2+1=0$

Example 2

- $(x-1)(x^2+1)=0$

$\equiv$ $\{$zero product rule: $ab=0 \equiv a=0 \vee b=0\}$

$x-1=0 \vee x^2+1=0$

$\equiv$ $\{$add 1 to both sides in left disjunct$\}$

$x=1 \vee x^2+1=0$

$\equiv$ $\{$add -1 to both sides in right disjunct$\}$

$x=1 \vee x^2=-1$

Example 2

- $(x-1)(x^2+1)=0$

$\equiv$ $\{$zero product rule: $ab=0 \equiv a=0 \vee b=0\}$

$x-1=0 \vee x^2+1=0$

$\equiv$ $\{$add 1 to both sides in left disjunct$\}$

$x=1 \vee x^2+1=0$

$\equiv$ $\{$add -1 to both sides in right disjunct$\}$

$x=1 \vee x^2=-1$

$\equiv$ $\{$a square is never negative$\}$

$x=1 \vee F$

Example 2

- $(x-1)(x^2+1)=0$

$\equiv$ {zero product rule: $ab=0 \equiv a=0 \vee b=0$}

$x-1=0 \vee x^2+1=0$

$\equiv$ {add 1 to both sides in left disjunct}

$x=1 \vee x^2+1=0$

$\equiv$ {add -1 to both sides in right disjunct}

$x=1 \vee x^2=-1$

$\equiv$ {a square is never negative}

$x=1 \vee F$

$\equiv$ {disjunction rule}

$x=1$

Sub derivations

There is often a need to do an auxiliary calculation, side proof or check a hypothesis while working on the main proof. This can be carried out in a nested derivation (a sub derivation).

Example 3: Determine the values of x for which the expression $\frac{x-1}{x^2-2}$ is well-defined.

Example 3

- $\frac{x-1}{x^2-1}$ is well-defined

Example 3

- $\frac{x-1}{x^2-1}$ is well-defined

$\equiv$ {definedness of rational expressions}

$x^2 - 1 \neq 0$

Example 3

- $\frac{x-1}{x^2-1}$ is well-defined

$\equiv$ {definedness of rational expressions}

$x^2 - 1 \neq 0$

$\equiv$ {switch to logic notation}

$\neg(x^2 - 1 = 0)$

Example 3

- $\frac{x-1}{x^2-1}$ is well-defined

$\equiv$ {definedness of rational expressions}

$x^2 - 1 \neq 0$

$\equiv$ {switch to logic notation}

$\neg(x^2 - 1 = 0)$

$\equiv$ {solve equation in brackets}

- $x^2 - 1 = 0$

  $\equiv$ {factorization rule}

  $(x+1)(x-1) = 0$

Example 3

- $\frac{x-1}{x^2-1}$ is well-defined

$\equiv$ {definedness of rational expressions}

$x^2 - 1 \neq 0$

$\equiv$ {switch to logic notation}

$\neg(x^2 - 1 = 0)$

$\equiv$ {solve equation in brackets}

- $x^2 - 1 = 0$

  $\equiv$ {factorization rule}

  $(x+1)(x-1) = 0$

  $\equiv$ {rule for zero product}

  $x = -1 \vee x = 1$

Example 3

- $\frac{x-1}{x^2-1}$ is well-defined

$\equiv$ {definedness of rational expressions}

$x^2 - 1 \neq 0$

$\equiv$ {switch to logic notation}

$\neg(x^2 - 1 = 0)$

$\equiv$ {solve equation in brackets}

- $x^2 - 1 = 0$

  $\equiv$ {factorization rule}

  $(x+1)(x-1) = 0$

  $\equiv$ {rule for zero product}

  $x = -1 \vee x = 1$

... $\neg(x = -1 \vee x = 1)$

$$\neg(x = -1 \vee x = 1)$$

$\equiv$ {de Morgans laws}

$$\neg(x = -1) \wedge \neg(x = 1)$$

$$\neg(x = -1 \vee x = 1)$$

$\equiv$ {de Morgans laws}

$$\neg(x = -1) \wedge \neg(x = 1)$$

$\equiv$ {change notation}

$$x \neq -1 \wedge x \neq 1$$

$$\neg(x = -1 \vee x = 1)$$

$\equiv$ {de Morgans laws}

$$\neg(x = -1) \wedge \neg(x = 1)$$

$\equiv$ {change notation}

$$x \neq -1 \wedge x \neq 1$$

$\square$

Hiding the sub derivation

$\bullet \quad \frac{x-1}{x^2-1}$ is well-defined

$\equiv \quad$ {definedness of rational expressions}

$\quad x^2 - 1 \neq 0$

$\equiv \quad$ {switch to logic notation}

$\quad \neg(x^2 - 1 = 0)$

$\equiv \quad$ {solve equation in brackets}

$\ldots \quad \neg(x = -1 \vee x = 1)$

$\equiv \quad$ {de Morgans laws}

$\quad \neg(x = -1) \wedge \neg(x = 1)$

$\equiv \quad$ {change notation}

$\quad x \neq -1 \wedge x \neq 1$

Explicit task and assumptions

The previous derivations are essentially just proofs:

- the task to be solved is not stated explicitly,
- nor are the assumptions that the proof may rely on listed explicitly.

The following example illustrates explicit tasks and assumptions in structured derivations.

Example 4: We show that if a, b and c are non-negative real numbers, then

$$(1+a)(1+b)(1+c) \geq 1+a+b+c$$

Example 4

- Show that $(1+a)(1+b)(1+c) \geq 1+a+b+c$

- when $a, b, c \geq 0$

Example 4

- Show that $(1+a)(1+b)(1+c) \geq 1+a+b+c$

- when $a,b,c \geq 0$

$\Vdash$

$$(1+a)(1+b)(1+c)$$

Example 4

- Show that $(1+a)(1+b)(1+c) \geq 1+a+b+c$

\- when $a,b,c \geq 0$

$\Vdash$

$(1+a)(1+b)(1+c)$

$=$ {multiply two last parenthesis}

$(1+a)(1+b+c+bc)$

Example 4

- Show that $(1+a)(1+b)(1+c) \geq 1+a+b+c$

\- when $a, b, c \geq 0$

$\Vdash$

$(1+a)(1+b)(1+c)$

$=$ {multiply two last parenthesis}

$(1+a)(1+b+c+bc)$

$=$ {multiply remaining parenthesis}

$1+b+c+bc+a+ab+ac+abc$

Example 4

- Show that $(1+a)(1+b)(1+c) \geq 1+a+b+c$

- when $a, b, c \geq 0$

$\Vdash$

$(1+a)(1+b)(1+c)$

$=$ {multiply two last parenthesis}

$(1+a)(1+b+c+bc)$

$=$ {multiply remaining parenthesis}

$1+b+c+bc+a+ab+ac+abc$

$\geq$ {expression *ab*, *ac*, *bc*, and *abc* are non-negative by assumption, subtract these}

$1+a+b+c$

Example 4

- Show that $(1+a)(1+b)(1+c) \geq 1+a+b+c$

\- when $a,b,c \geq 0$

$\Vdash$ {combining $=$ and $\geq$ gives $\geq$}

$(1+a)(1+b)(1+c)$

$=$ {multiply two last parenthesis}

$(1+a)(1+b+c+bc)$

$=$ {multiply remaining parenthesis}

$1+b+c+bc+a+ab+ac+abc$

$\geq$ {expression ab, ac, bc, and abc are non-negative by assumption, subtract these}

$1+a+b+c$

Example 4

- Show that $(1+a)(1+b)(1+c) \geq 1+a+b+c$

\- when $a,b,c \geq 0$

$\Vdash$ {combining $=$ and $\geq$ gives $\geq$}

$(1+a)(1+b)(1+c)$

$=$ {multiply two last parenthesis}

$(1+a)(1+b+c+bc)$

$=$ {multiply remaining parenthesis}

$1+b+c+bc+a+ab+ac+abc$

$\geq$ {expression *ab*, *ac*, *bc*, and *abc* are non-negative by assumption, subtract these}

$1+a+b+c$

□

Informal reasoning in structured derivations

Structured derivations can be carried out at any level of detail in the proof:

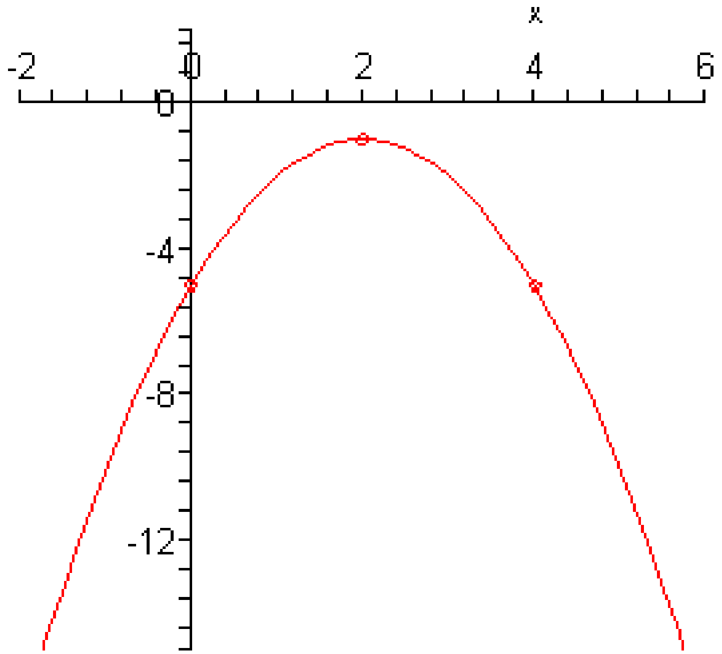
- from very detailed, even axiomatic proofs
- to high level proofs that mainly outline the argumentation

Example 5: Determine the values of constant a such that the function

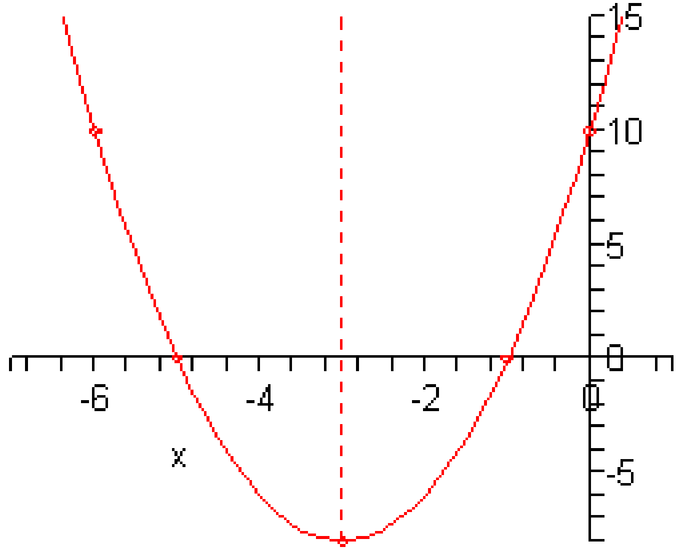
$$f(x) = -x^2 + ax + a - 3$$

is always negative.

Parabolas



Parabola with negative coefficient for the squared term



Parabola with positive coefficient for the squared term

Example 5

- Determine the values of constant a such that the function $f(x) = -x^2 + ax + a - 3$ is always negative.:

$\Vdash$ $(\forall x \cdot -x^2 + ax + a - 3 < 0)$

Example 5

- Determine the values of constant a such that the function $f(x) = -x^2 + ax + a - 3$ is always negative.:

$\Vdash$ $(\forall x \cdot -x^2 + ax + a - 3 < 0)$

$\equiv$ {the function is a parabola that opens downwards, as the coefficient for x^2 is negative; such a function is always negative if it does not intersect the x- axis (figure above on the left)}

$(\forall x \cdot -x^2 + ax + a - 3 \neq 0)$

Example 5

- Determine the values of constant a such that the function $f(x) = -x^2 + ax + a - 3$ is always negative.:

$\Vdash$ $(\forall x \cdot -x^2 + ax + a - 3 < 0)$

$\equiv$ {the function is a parabola that opens downwards, as the coefficient for x^2 is negative; such a function is always negative if it does not intersect the x- axis (figure on the left)}

$(\forall x \cdot -x^2 + ax + a - 3 \neq 0)$

$\equiv$ {this is equivalent to the discriminant D for the equation being negative}

$D < 0$

$\equiv$ {compute D}

- D

$\equiv$ {compute D}

$\bullet$ D

$=$ {the discriminant for the equation $Ax^2 + Bx + C = 0$ is $B^2 - 4AC$}

$a^2 - 4(-1)(a-3)$

$\equiv$ {compute D}

$\bullet$ D

$=$ {the discriminant for the equation $Ax^2 + Bx + C = 0$ is $B^2 - 4AC$}

$a^2 - 4(-1)(a-3)$

$=$ {simplify}

$a^2 + 4a - 12$

$\equiv$ {compute D}

- D

$=$ {the discriminant for the equation $Ax^2 + Bx + C = 0$ is $B^2 - 4AC$}

$a^2 - 4(-1)(a-3)$

$=$ {simplify}

$a^2 + 4a - 12$

$\ldots$ $a^2 + 4a - 12 < 0$

$\equiv$ {the function $a^2 + 4a - 12$ is a parabola that opens upwards, because the coefficient for a^2 is positive; such a function is negative between the intersection points with the x- axis (figure above on right)}

$\equiv$ {the function $a^2+4a-12$ is a parabola that opens upwards, because the coefficient for a^2 is positive; such a function is negative between the intersection points with the x- axis (figure on right)}

- Compute the places where $a^2+4a-12$ intersects the x - axis:

$\Vdash$ $a^2+4a-12=0$

$\equiv$ {the function $a^2+4a-12$ is a parabola that opens upwards, because the coefficient for a^2 is positive; such a function is negative between the intersection points with the x- axis (figure on right)}

- Compute the places where $a^2+4a-12$ intersects the x - axis:

$\Vdash$ $a^2+4a-12=0$

$\equiv$ {square root formula}

$$a=\frac{-4\pm\sqrt{4^2-4\cdot1\cdot(-12)}}{2\cdot1}$$

≡ {the function $a^2+4a-12$ is a parabola that opens upwards, because the coefficient for a^2 is positive; such a function is negative between the intersection points with the x- axis (figure on right)}

- Compute the places where $a^2+4a-12$ intersects the x - axis:

⊩ $a^2+4a-12=0$

≡ {square root formula}

$a=\frac{-4\pm\sqrt{4^2-4\cdot1\cdot(-12)}}{2\cdot1}$

≡ {simplify}

$a=2\vee a=-6$

$\equiv$ {the function $a^2+4a-12$ is a parabola that opens upwards, because the coefficient for a^2 is positive; such a function is negative between the intersection points with the x- axis (figure on right)}

- Compute the places where $a^2+4a-12$ intersects the x - axis:

$\Vdash$ $a^2+4a-12=0$

$\equiv$ {square root formula}

$a=\frac{-4\pm\sqrt{4^2-4\cdot 1\cdot(-12)}}{2\cdot 1}$

$\equiv$ {simplify}

$a=2\vee a=-6$

$\ldots$ $-6<a<2$

$\square$

Complete derivation

- Determine the values of constant a such that the function $f(x) = -x^2 + ax + a - 3$ is always negative.:

$\Vdash$ $(\forall x \cdot -x^2 + ax + a - 3 < 0)$

$\equiv$ {the function is a parabola that opens downwards, as the coefficient for x^2 is negative; such a function is always negative if it does not intersect the x- axis (figure on the left)}

$(\forall x \cdot -x^2 + ax + a - 3 \neq 0)$

$\equiv$ {this is equivalent to the discriminant D for the equation being less than zero}

$D < 0$

$\equiv$ {substitute values for D}

  - Compute the discriminant D:

    D

    $=$ {the discriminant for the equation $Ax^2 + Bx + C = 0$ is $B^2 - 4AC$}

$$
\begin{array}{lll}
 & & a^2 - 4(-1)(a-3) \\
 & = & \{\text{simplify}\} \\
 & & a^2 + 4a - 12
\end{array}
$$

$\ldots \quad a^2 + 4a - 12 < 0$

$\equiv$ {the function $a^2 + 4a - 12$ is a parabola that opens upwards, because the coefficient for a^2 is positive; such a function is negative between the intersection points with the x- axis (figure on right)}

- Compute the places where $a^2 + 4a - 12$ intersects the x - axis:

$$
\begin{array}{ll}
\Vdash & a^2 + 4a - 12 = 0 \\
\equiv & \{\text{square root formula}\} \\
 & a = \frac{-4 \pm \sqrt{4^2 - 4 \cdot 1 \cdot (-12)}}{2 \cdot 1} \\
\equiv & \{\text{simplify}\} \\
 & a = 2 \vee a = -6
\end{array}
$$

$\ldots \quad -6 < a < 2$

$\square$

Hiding sub derivations

- hiding the sub derivations shows the overall structure and main argumentation of the proof
- an outlining editor that supports collapsing (possible in, e.g., Lyx and TexMacs - two wysiwyg latex editors) is very useful for writing and browsing structured derivations

- Determine the values of constant a such that the function $f(x) = -x^2 + ax + a - 3$ is always negative.:

$\Vdash$ $(\forall x \cdot -x^2 + ax + a - 3 < 0)$

$\equiv$ {the function is a parabola that opens downwards, as the coefficient for x^2 is negative; such a function is always negative if it does not intersect the x- axis (figure on the left)}

$(\forall x \cdot -x^2 + ax + a - 3 \neq 0)$

$\equiv$ {this is equivalent to the discriminant D for the equation being zero}

$D < 0$

$\equiv$ {substitute values for D}

... $a^2 + 4a - 12 < 0$

$\equiv$ {the function $a^2 + 4a - 12$ is a parabola that opens upwards, because the coefficient for a^2 is positive; such a function is negative between the intersection points with the x- axis (figure on right)}

... $-6 < a < 2$

□

Using observations in derivations

A common way of solving problems:

- Write down the task to be solved and the assumptions that we are allowed to make
- Write down some simple observations that can be directly made, based on the assumptions
- Continue with some additional observations based on previous observations
- Derive a solution to the original problem based on these observations.

Example: percentages

Example 6. The IOP party got 20 % of the votes last year. This year it got 26 % of the votes. How many percentages did the relative share of votes grow.

(Original: Oppilaskunnan vaaleissa IOP sai 20 % äänistä viime vuonna. Tänä vuonna 26% äänistä. Kuinka monta prosenttia suhteellinen osuus äänistä kasvoi.)

Example 6

- Compute the increase in relative vote share for IOP i(n percentage), when

- original share of votes for IOP = 20%

- new share of votes for IOP = 26%

Example 6

- Compute the increase in relative vote share for IOP i(n percentage), when

\- original share of votes for IOP = 20%

\- new share of votes for IOP = 26%

\+ {increase in share of votes = new share of votes - original share of votes}

increase in share of votes = 26% - 20% = 6%

Example 6

- Compute the increase in relative vote share for IOP i(n percentage), when

- original share of votes for IOP = 20%

- new share of votes for IOP = 26%

+ {increase in share of votes = new share of votes - original share of votes}

  increase in share of votes = 26% - 20% = 6%

+ {increase in relative share of votes = increase in share of votes/ original share of votes}

  increase in relative share of votes = 6 % / 20 % = 0.3

Example 6

- Compute the increase in relative vote share for IOP i(n percentage), when

\- original share of votes for IOP = 20%

\- new share of votes for IOP = 26%

\+ {increase in share of votes = new share of votes - original share of votes}

increase in share of votes = 26% - 20% = 6%

\+ {increase in relative share of votes = increase in share of votes/ original share of votes}

increase in relative share of votes = 6 % / 20 % = 0.3

⊩ {observations above}

increase in relative vote share for IOP (in percentage)

Example 6

- Compute the increase in relative vote share for IOP i(n percentage), when

\- original share of votes for IOP = 20%

\- new share of votes for IOP = 26%

\+ {increase in share of votes = new share of votes - original share of votes}

increase in share of votes = 26% - 20% = 6%

\+ {increase in relative share of votes = increase in share of votes/ original share of votes}

increase in relative share of votes = 6 % / 20 % = 0.3

$\Vdash$ {observations above}

increase in relative vote share for IOP (in percentage)

$=$ {increase in relative share (in percentage) = increase in relative share of votes $\times$100 %}

0.3 $\times$100 %

Example 6

- Compute the increase in relative vote share for IOP i(n percentage), when

\- original share of votes for IOP = 20%

\- new share of votes for IOP = 26%

\+ {increase in share of votes = new share of votes - original share of votes}

increase in share of votes = 26% - 20% = 6%

\+ {increase in relative share of votes = increase in share of votes/ original share of votes}

increase in relative share of votes = 6 % / 20 % = 0.3

$\Vdash$ {observations above}

increase in relative vote share for IOP (in percentage)

$=$ {increase in relative share (in percentage) = increase in relative share of votes $\times$100 %}

0.3 $\times$100 %

$=$ {calculate}

30%

$\square$

Esimerkki: prosenttilaskenta

- Laske kuinka monta prosenttia suhteellinen osuus äänistä kasvoi, kun

\- alkuperäinen ääniosuus = 20%

\- uusi ääniosuus = 26%

\+ {ääniosuuden kasvu = uusi ääniosuus - alkuperäinen ääniosuus}

ääniosuuden kasvu = 26% - 20% = 6%

\+ {suhteellinen osuuden kasvu = ääniosuuden kasvu/ alkuperäinen ääniosuus}

suhteellinen osuuden kasvu = 6 % / 20 % = 0.3

\+ {suhteellinen osuuden kasvuprosentti = suhteellinen osuuden kasvu $\times$ 100 %}

suhteellinen osuuden kasvuprosentti = 0.3 $\times$ 100 % = 30%

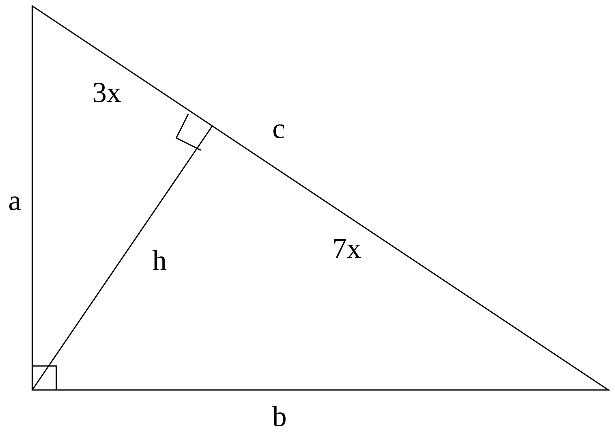
\+ {oletusten tarkkuus on yksi merkitsevä numero, niin vastaus annetaan yhden numeron tarkkuudella}

$\Vdash$ {yo laskeman mukaan suhteellinen ääniosuuden kasvuprosentti = 30 %}

Using observations in structured derivations

In many proofs, it is useful to make some preliminary observations, before one tackles the main proof. This is particularly useful in forward proofs (common in, e.g., geometry):

Example 7: Consider a right triangle with catheters a and b and hypotenuse c. Assume that the height of the triangle on the hypotenuse divides the hypotenuse in the proportion 3:7. Determine the proportion $\frac{a}{b}$.

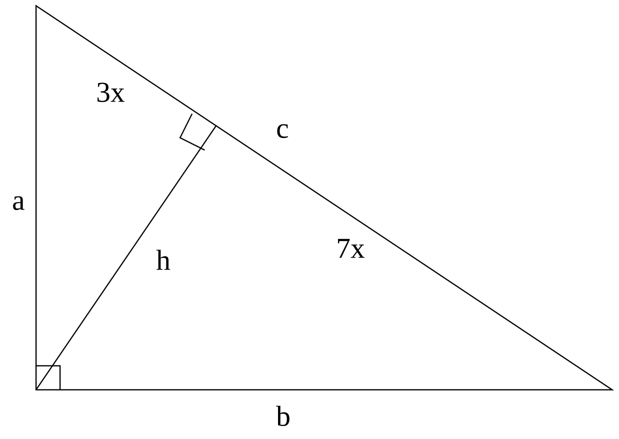


Example 7

- Determine $\frac{a}{b}$, when

- the triangle is right, with hypotenuse c and catheters a and b,

- the height of the triangle on the hypotenuse divides the hypotenuse in the proportion 3:7



[1] {from figure}

$c = 10x$

[2] {figure and Pythagorean theorem}

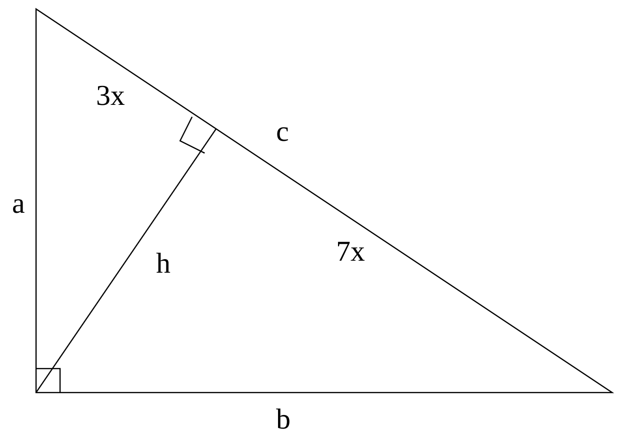
$h^2 + 9x^2 = a^2$

[3] {figure and Pythagorean theorem}

$h^2 + 49x^2 = b^2$

[4] {figure, observation [1] and Pythagorean theorem}

$a^2 + b^2 = 100x^2$



[5] {subtract equation [2] from equation [3] and simplify}

$b^2 - a^2 = 40x^2$

[6] {add equations [4] and [5] and simplify}

$b^2 = 70x^2$

[7] {substitute equation [6] into equation [5]}

- $a^2 + b^2 = 100x^2$

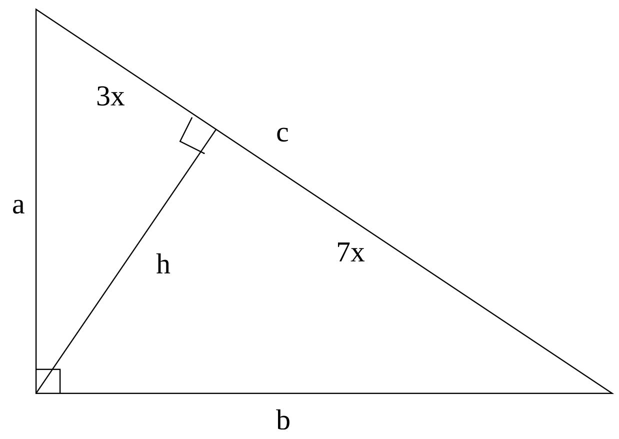
$\equiv$ {equation [6]}

$a^2 + 70x^2 = 100x^2$

$\equiv$ {solve a^2}

$a^2 = 30x^2$

... $a^2 = 30x^2$



$\Vdash$ $\frac{a}{b}$

$=$ {square root definition, a and b are positive numbers}

$\sqrt{\frac{a^2}{b^2}}$

$=$ {observations [6] and [7]}

$\sqrt{\frac{30x^2}{70x^2}}$

$=$ {simplify}

$\frac{\sqrt{3}}{\sqrt{7}}$

□

Research on teaching structured derivations in high school

- Research carried out in *Learning and Reasoning laboratory* (Departments of Information Technology at Åbo Akademi University och University of Turku).
- Focus on methods for teaching mathematics and programming
- Experimental approach:
  - develop methods, notation, tools and teaching material for mathematics and program construction (formal methods research)
  - try out the methods in practical teaching experiments, measure results qualitatively and quantitatively
  - use the feedback from the experiments to improve the methods, tools, and teaching material
- We have a web-based resource center, *IMPEd* (crest.cs.abo.fi/imped/) with research and teaching material

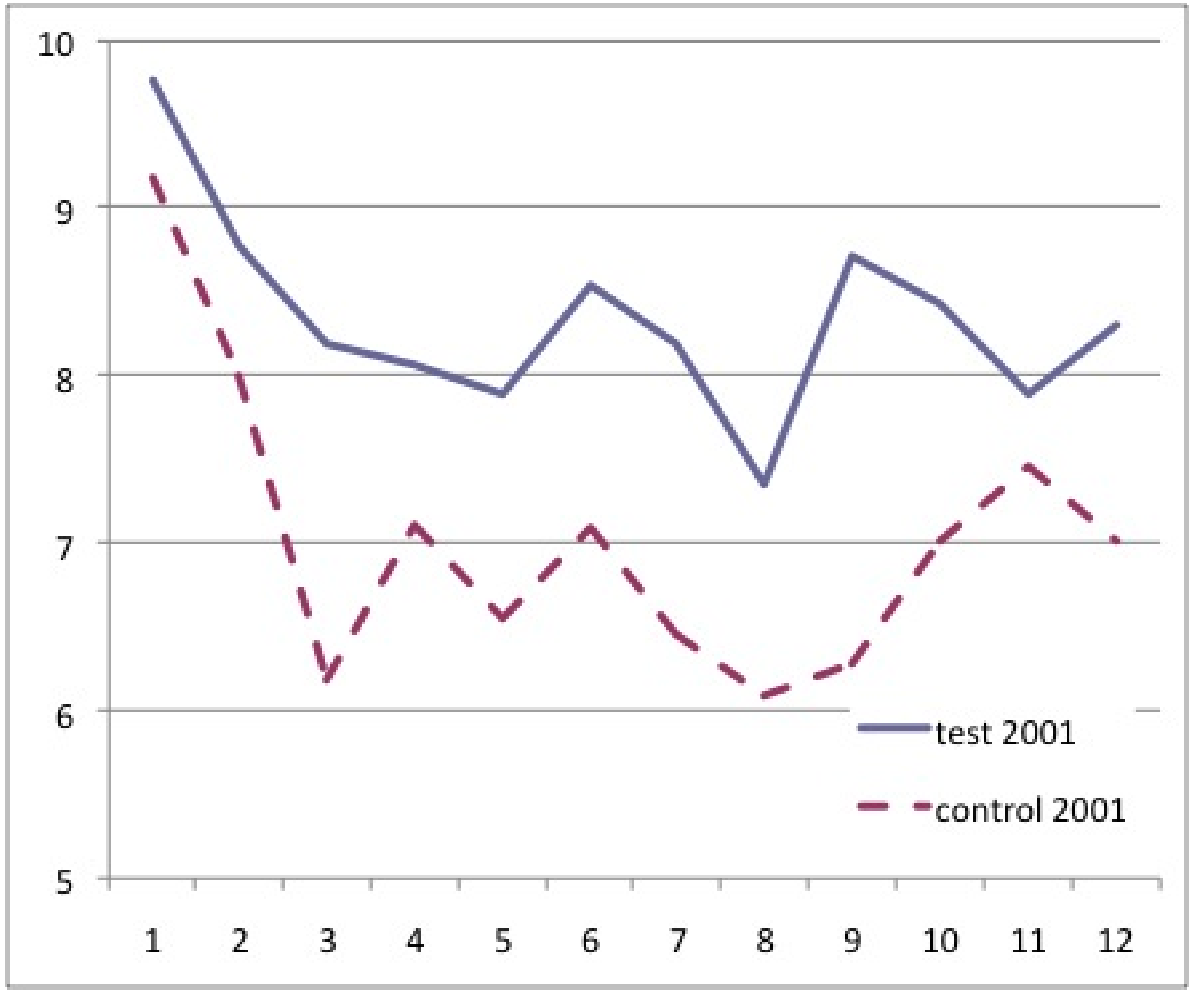
Empirical studies on using structured derivations in teaching

- Tried out structured derivations on a large (and unbiased) collection of high school math problems (Finnish national matriculation exams 1998 - 2009, 180 assignments all together, tests all math courses in high school).
- Two large comparative studies on using structured derivations throughout the high school math curriculum: one group of students were taught all math courses in high school using structured derivations, while a control group was taught the same courses in the standard way.
- Special elective course in high school ("Logic and number theory") that introduces structured derivations together with basic logic, and applies it to number theory. The course has been lectured appr. 10 times, in different high schools, with very promising results.
- Redesigned first year CS curriculum at the IT department at Åbo Akademi : start with a course on practical programming (Python) and logic (structured derivations), and then follow up with course on correct-by-design programming (invariant based programming).
- Teaching experiment in Turku Polytechnic in autumn 2008: a course in trigonometry that was taught using structured derivations

Kupittaa high school experiment (Peltomäki)

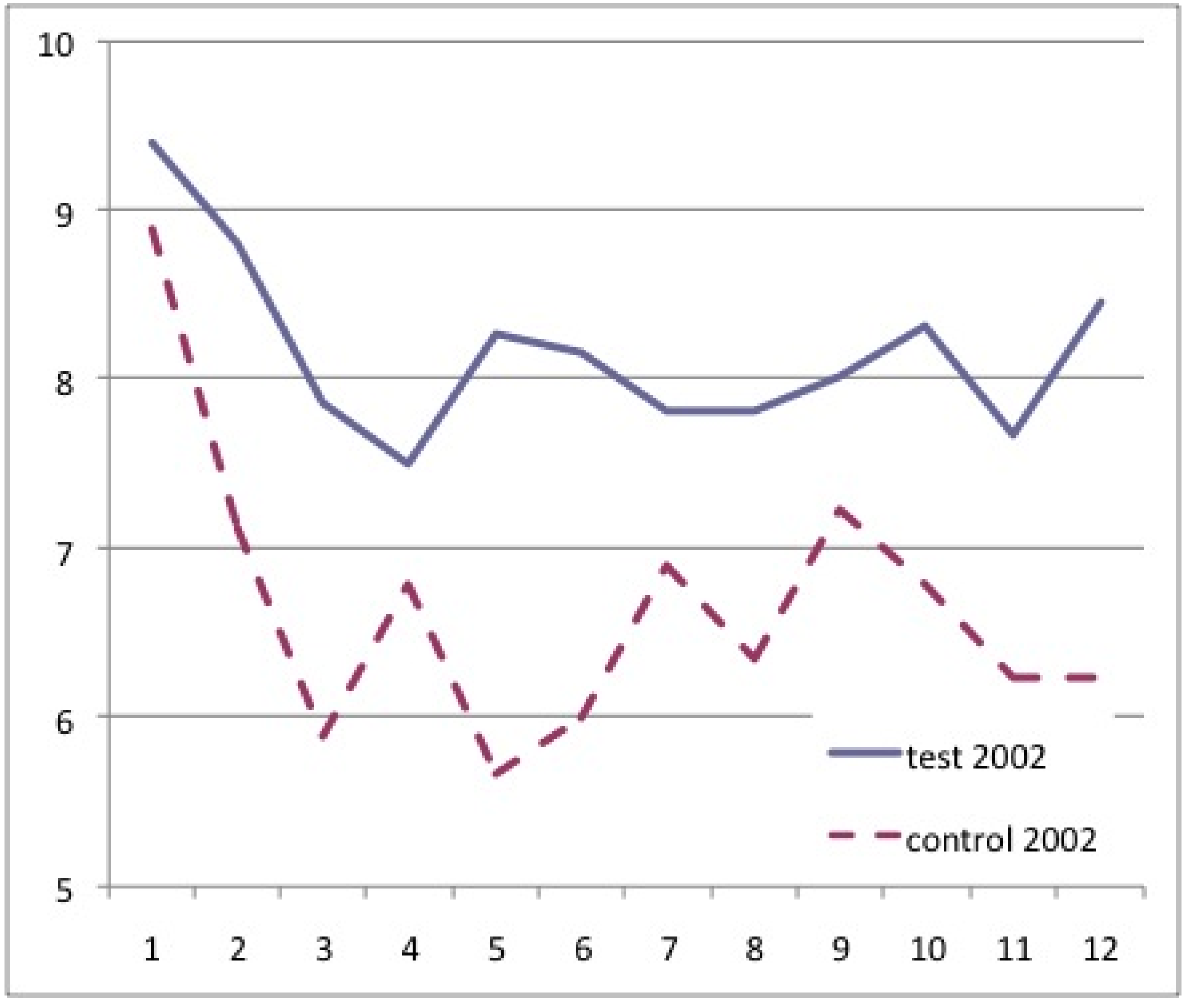
- High school math in Finland is given on two levels, advanced (taken by 40% of students) and normal (60% of students). Advanced level required for Engineering, Science, Medicine and Business studies at university.
- Two successive experiments, 2001 - 2004 and 2002 - 2005 at advanced level math
- New students at Kupittaa high school (Turku) were divided into three groups:
  - a test group consisting of those who choose IT as specialization and who were taught math using structured derivations
  - a control group was taught math in the standard way
  - a third group that did not participate in experiment
- The control group was chosen to be as similar as possible to the test group (but final choice of which groups to join was free for students)
- The test and control groups had same course content, same assignments and exams, lectures and exams at the same time, but different teachers.

Results 2001 - 2004, class average grade



1 junior high math grade
2 functions and equations 1
3 functions and equations 2
4 geometry
5 trigonometry and vectors
6 analytic geometry
7 differential equations 1
8 differential equations 2
9 integrals
10 statistics and prob. theory
11 sequences and series
12 matriculation exam

Results 2002 - 2005, class average grade



1 junior high math grade
2 functions and equations 1
3 functions and equations 2
4 geometry
5 trigonometry and vectors
6 analytic geometry
7 differential equations 1
8 differential equations 2
9 integrals
10 statistics and prob. theory
11 sequences and series
12 matriculation exam

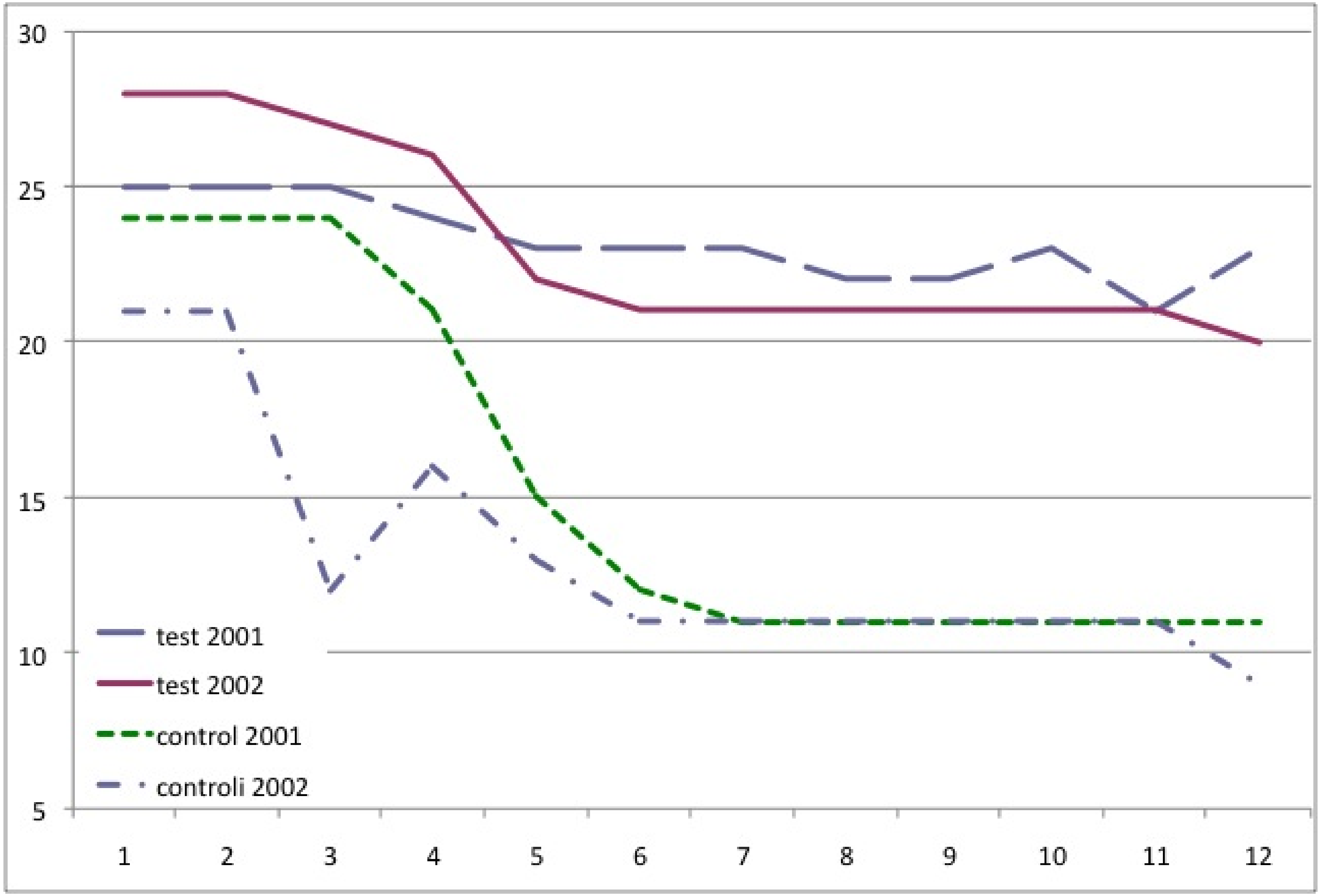
Matriculation exam 2004 grades

	test group	control group	whole school	whole country
N	21	11	50	12 494
laudatur	24%	0%	10%	6%
eximia	43%	9%	22%	15%
laudatur+eximia	67%	9%	32%	21%
average (5 - 10)	8.48	7.00	7.24	6.92

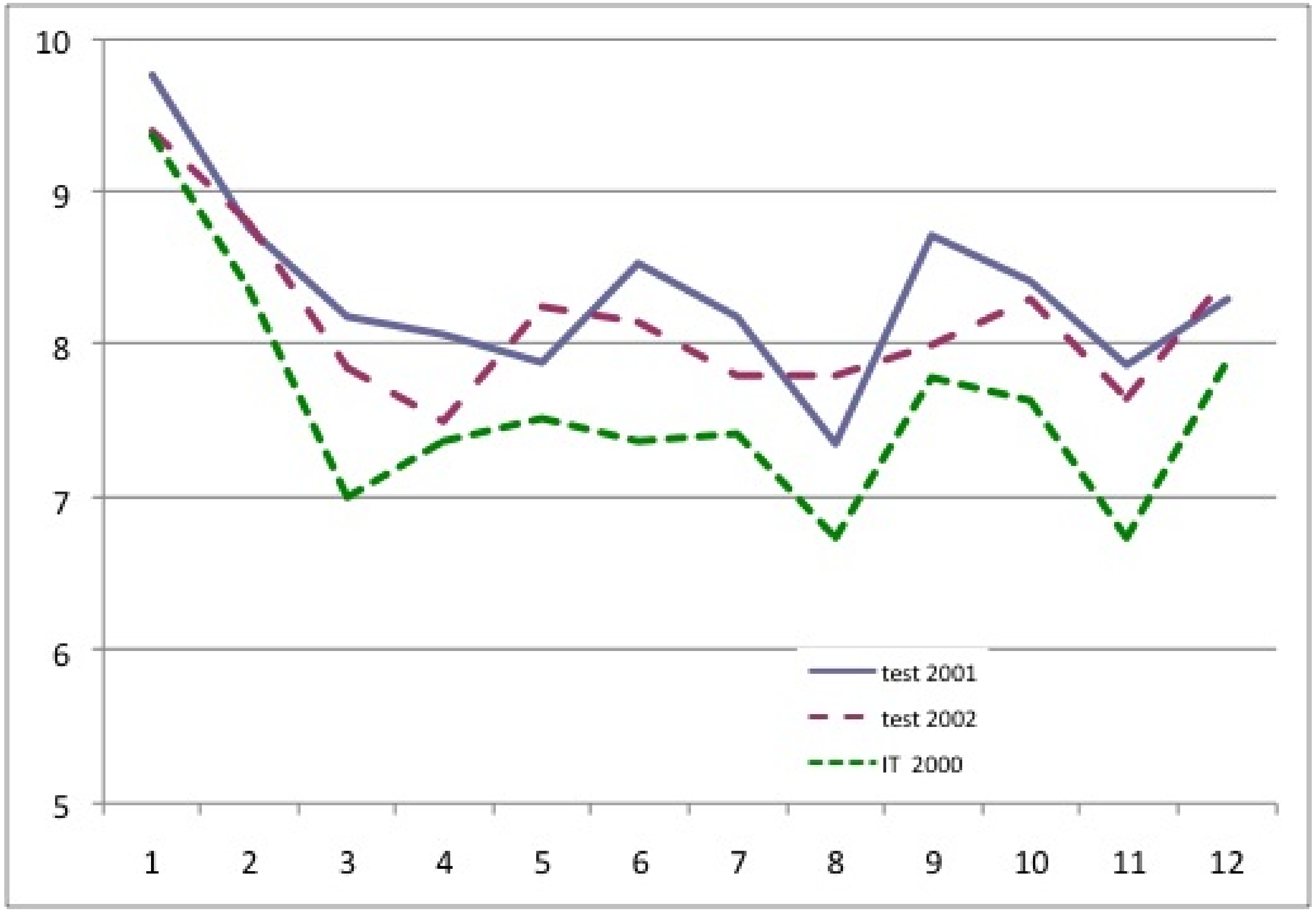
Matriculation exam 2005 grades

	test group	control group	whole school	whole country
N	20	9	40	11 627
laudatur	10%	0%	4%	6%
eximia	50%	0%	20%	16%
laudatur+eximia	60%	0%	24%	22%
average (5 - 10)	8.45	6.22	7.00	6.96

Interrupted in both studies



Comparison with previous year IT-group



- same selection criteria
- same teacher
- same junior high degree as test2002

Summary of Kupittaa experiments

- Test group performed clearly better than the control group
- Matriculation exams in particular much better for test group
- Much lower levels of interrupted studies for test group
- The test groups had in both studies somewhat better junior high school degree in math compared to the control group. The teacher for the test group has also a good track record.
- Difference in favor of test group persist even if these differences are eliminated

Ongoing and planned activities

- Finnish National Board of Education has commissioned a number of courses on structured derivations for high school teachers (continuous training courses, 3 -5 days) in 2008 - 2009
- Teaching experiments in lower secondary education (junior high school) are planned in a some schools in the Turku region
- A number of training courses for junior high school teachers are planned for 2010
- City of Turku is planning to expand the use of structured derivations to all high schools in the region
- The Technology Industry Foundation of Finland is funding our resource center for structured derivations (Imped)
- The Technology Development Fund of Finland (TEKES) is financing a research project on using computers and structured derivations for teaching math at junior high level.

Imped resource center (crest.cs.abo.fi/imped)

- Research publications
- Talks and presentations
- Tool support (Lyx editor adapted to structured derivations)
- Courses and course material, support for web based courses
- Tutorials on using structured derivations in teaching mathematics (new lecture notes series)
- Published 6 tutorials (presently only available in Swedish and Finnish)

Tool support for structured derivations

- Wysiwyg latex editors: Lyx
  - support full mathematical formalism, see almost final text when editing
  - support for outlining (hiding and showing sub derivations)
- Adapted LyX to structured derivations
- Plan is to adapt it also to a learning environment (Moodle)
- Automatic checking of structured derivations
  - syntax checking of structured derivations
  - interface to PVS and Simplify
  - proof checker checks each individual step (under construction)
- Checking correctness requires more formalized language than what is common in high school math

Thank you for listening !

Structured derivations in CS education

- Logic course given since 1996 as introductory CS course (based on Gries' and Schneider's book)
- Changed the course to be based on structured derivations in 2006 (same as high school course on "Logic and number theory")
- Course went otherwise well, students learned the method, but high school material was considered to be too simple
- Course was given anew in 2007, but reworked to make the math more challenging (lattice theory, elementary algebra, discrete math) and include more systematic presentation of logic (propositional calculus, predicate calculus)
- This now worked well with the students, they liked the course, results were good, and feedback from students was very positive
- Course now standard and compulsory for first year students.

Structured derivations as a unified proof format

Syntax for structured derivations

The syntax is recursive:

- Derivations are defined in terms of justifications
- Justifications are defined in terms of derivations

As a consequence, structured derivations can be arbitrarily deeply nested.

Intuitive syntax

The syntax should be

- as simple as possible
- as intuitive as possible
- easy to remember and apply
- as few additional symbols as possible
- allow for common shortcuts and abbreviations
- not force a verbose expression where a simple one would do

The intuitive syntax is shown below, the formal is described later

derivation:		**justification**:
•	task	{motivation}
-	assumption	derivation
	⋮	derivation
+	justification	⋮
	observation	
	⋮	
$\Vdash$	justification	
	term	
rel	justification	
	term	
	⋮	
□		

General form for structured derivations

A derivation consists of four parts:

- The task (problem): what shall we do, what assumptions can we make
- Observations: what facts can we infer directly from the assumptions and previous observations
- Justification for the solution of the task: why do the observations and the calculation together solve the original problem
- Calculation: computing the required result

derivation:

•	task	the problem
-	assumptions	
	⋮	
+	justification	justifications
	observation	
	⋮	
⊩	justification	justifying the solution
	term	calculation
rel	justification	
	term	
	⋮	
□	. . . .	the proof is complete

General proof format

The different parts correspond to different proof formats:

- Observations: Hilbert-like proof
- Justifying the solution: Gentzen-like proof
- Calculation: Dijkstra-like proof

derivation:

- • task . . . . the problem
- \- assumptions

  ⋮

- \+ justification . . . . Hilbert-like proof

  observation

  ⋮

- ⊩ justification . . . . Gentzen-like proof

  term . . . . Dijkstra-like proof

- rel justification

  term

  ⋮

- □ . . . . the proof is complete

Common proof formats

The standard proof formats result from leaving out some components of structured derivations

- Omitting observations, the problem (task and assumptions), sub derivations and the justification for the solution gives us Dijkstra's original calculational style proofs.
- Omitting the calculation and sub derivations gives us Hilbert-like proofs.
- Omitting observations and calculations gives us Gentzen-like proofs.

Mixing proof formats

Structured derivations allow all these different proof formats to be freely mixed in a single derivation. We are then free to choose the specific proof method that is appropriate for the specific (sub)problem considered.

The mixing is achieved by

1. the fact that a single structured derivation can contain a Hilbert-like part, a Gentzen-like part and a Dijkstra-like part

2. a sub derivation can use a different proof format from the main derivation, or a combination of proof formats.

Dijkstra-like

- term

= {motivation}

term

⋮

□

Hilbert-like

- task

\- assumptions

⋮

\+ {motivation}

observations

⋮

⊩ {motivation}

□

Gentzen-like

- task

\- assumption

⋮

⊩ justification

□

Formal syntax definition

Formal syntax is needed for computer support: checking correctness of syntax, analysing correctness of derivation steps, annotating the derivation, etc.

The syntax is defined on four different levels:

- structure of structured derivations
- layout for structured derivations
- parameters
- dialects

Structure

$$
\begin{aligned}
derivation &::= derivationId\, task \\
&\quad (assumptionId\, assumption)^* \\
&\quad (observationId\; justification\, observation)^* \\
&\quad startproof\; proof\; [endproof] \\
derivation &::= derivationId\, calculation\, [endproof] \\
proof &::= justification \mid calculation \mid justification\, calculation \\
calculation &::= term\, (relation\; justification\, term)^+ \\
justification &::= motivation\, [startsub\, derivation^+\; endsub]
\end{aligned}
$$

Layout

$$
\begin{aligned}
derivationId &::= \text{"}\bullet\text{"} \mid label\text{" : "} \\
assumptionId &::= \text{"}-\text{"} \mid \text{"("}label\text{")"} \\
observationId &::= \text{"}+\text{"} \mid \text{"["}label\text{"]"} \\
task &::= \text{tab}\, tasktText\, \text{ret} \\
assumption &::= \text{tab}\, assumptionText\, \text{ret} \\
observation &::= \text{tab}\, observationText\, \text{ret} \\
term &::= \text{tab}\, termtText\, \text{ret} \\
relation &::= relationText \\
motivation &::= \text{tab"\{"}motivationtText\text{"\}"ret} \\
startproof &::= \text{"}\Vdash\text{"} \mid \text{"}\Vdash_0\text{"} \\
endproof &::= \text{"}\square\text{"ret} \\
startsub &::= \text{indent} \\
endsub &::= \text{dedent"}\ldots\text{"}
\end{aligned}
$$

Parameters

$$
\begin{aligned}
\mathit{label} &::= \ \ldots \text{ notation for labels } \ldots \\
\mathit{taskText} &::= \ \ldots \text{ notation for tasks } \ldots \\
\mathit{assumptionText} &::= \ \ldots \text{ notation for assumptions } \ldots \\
\mathit{observationText} &::= \ \ldots \text{ notation for observations } \ldots \\
\mathit{termText} &::= \ \ldots \text{ notation for terms } \ldots \\
\mathit{relationText} &::= \ \ldots \text{ notation for relations} \ldots \\
\mathit{motivationText} &::= \ \ldots \text{ notation for motivations } \ldots
\end{aligned}
$$

Dialects

- Informal structured derivations
- Structured derivations with logic
- Axiomatic proofs
- Structured derivations for proof checkers

Informal dialect

We use structured derivations informally, using whatever ad hoc notation feels most suitable for the problem at hand.

In this case, we allow any text string for describing the task, the assumptions, the observations, the terms and the motivations.

Labels are used in a traditional way, and are usually Arabic or roman numerals, or alphabetic characters.

This format is very loose, it does not even require any systematic use of logical notation in the derivations, and the motivations can be very informal and intuitive.

Logic dialect

$$
\begin{aligned}
label &::= \ldots \text{ notation for labels } \ldots \\
taskText &::= imperative\ proposition \\
assumptionText &::= proposition\ correction \\
observationText &::= proposition \\
termText &::= expression \\
relationText &::= relationSymbol \\
motivationText &::= \ldots \text{ notation for motivations } \ldots \\
imperative &::= \text{... expression for what needs to be done ...} \\
correction &::= (\text{"\#\#"}\ proposition)^*
\end{aligned}
$$

Axiomatic dialect

Logic + following definitions

$$\begin{aligned} label &::= number \\ motivationText &::= ruleName\,[rule]\,assignment\,validity \\ imperative &::= \text{""} \end{aligned}$$

Proof checker dialect

Logic + following definition

$$motivationText \ ::= \ strategy$$