Online proof and elementary mathematics from an educational perspective

Chris Sangwin

School of Mathematics
University of Edinburgh

May 2019

Research question

To what extent can students' mathematical proofs be automatically assessed?

Research question

To what extent can students' mathematical proofs be automatically assessed?

Focus: the proof and reasoning which occurs in current mathematics examinations.

Research question

To what extent can students' mathematical proofs be automatically assessed?

Focus: the proof and reasoning which occurs in current mathematics examinations.

Corpus: 2018 paper from the SQA Advanced Higher Mathematics examinations.
($\cdots \equiv$ Further Mathematics $\equiv$ IB HL)

Caveat....

This is a long way from the Kepler Conjecture....

Caveat....

This is a long way from the Kepler Conjecture....
... but a lot more people learn this mathematics ...

Caveat....

This is a long way from the Kepler Conjecture....
... but a lot more people learn this mathematics ...
... and almost everyone starts this way.

Automatic assessment with STACK

Differentiate $\frac{e^{5\,x}}{7\,x+1}$ with respect to x.

```
diff(e^(5*x)/(7*x+1),x)
= (5*(7*x+1)*e^(5*x)-7*e^(5*x))/(7*x+1)^2
= (35*x-2)*e^(5*x)/(7*x+1)^2
```

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{e^{5\,x}}{7\,x+1} \qquad x \notin \left\{-\frac{1}{7}\right\}$$

$$\checkmark \quad = \frac{5\,(7\,x+1)\,e^{5\,x}-7\,e^{5\,x}}{(7\,x+1)^2} \qquad x \notin \left\{-\frac{1}{7}\right\}$$

$$\checkmark \quad = \frac{(35\,x-2)\,e^{5\,x}}{(7\,x+1)^2} \qquad x \notin \left\{-\frac{1}{7}\right\}$$

The variables found in your answer were: $[x]$

Automatic assessment with STACK

Differentiate $\frac{e^{5\,x}}{7\,x+1}$ with respect to x.

```
diff(e^(5*x)/(7*x+1),x)
= (5*(7*x+1)*e^(5*x)-7*e^(5*x))/(7*x+1)^2
= (35*x-2)*e^(5*x)/(7*x+1)^2
```

$$
\begin{array}{cll}
 & \frac{d}{dx}\frac{e^{5\,x}}{7\,x+1} & x \notin \left\{-\frac{1}{7}\right\} \\
\checkmark & = \frac{5\,(7\,x+1)\,e^{5\,x}-7\,e^{5\,x}}{(7\,x+1)^2} & x \notin \left\{-\frac{1}{7}\right\} \\
\checkmark & = \frac{(35\,x-2)\,e^{5\,x}}{(7\,x+1)^2} & x \notin \left\{-\frac{1}{7}\right\}
\end{array}
$$

The variables found in your answer were: $[x]$

Notes:

- Students' answers have mathematical content

Automatic assessment with STACK

Differentiate $\frac{e^{5\,x}}{7\,x+1}$ with respect to x.

```
diff(e^(5*x)/(7*x+1),x)
= (5*(7*x+1)*e^(5*x)-7*e^(5*x))/(7*x+1)^2
= (35*x-2)*e^(5*x)/(7*x+1)^2
```

		$\frac{\mathrm{d}}{\mathrm{d}x}\frac{e^{5\,x}}{7\,x+1}$	$x \notin \left\{-\frac{1}{7}\right\}$
✓	$=$	$\frac{5\,(7\,x+1)\,e^{5\,x}-7\,e^{5\,x}}{(7\,x+1)^2}$	$x \notin \left\{-\frac{1}{7}\right\}$
✓	$=$	$\frac{(35\,x-2)\,e^{5\,x}}{(7\,x+1)^2}$	$x \notin \left\{-\frac{1}{7}\right\}$

The variables found in your answer were: $[x]$

Notes:

- Students' answers have mathematical content
- Current rudimentary interface for line by line working.

Automatic assessment with STACK

Differentiate $\frac{e^{5\,x}}{7\,x+1}$ with respect to x.

```
diff(e^(5*x)/(7*x+1),x)
= (5*(7*x+1)*e^(5*x)-7*e^(5*x))/(7*x+1)^2
= (35*x-2)*e^(5*x)/(7*x+1)^2
```

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{e^{5\,x}}{7\,x+1} \qquad x \notin \left\{-\frac{1}{7}\right\}$$

$$\checkmark \quad = \frac{5\,(7\,x+1)\,e^{5\,x}-7\,e^{5\,x}}{(7\,x+1)^2} \qquad x \notin \left\{-\frac{1}{7}\right\}$$

$$\checkmark \quad = \frac{(35\,x-2)\,e^{5\,x}}{(7\,x+1)^2} \qquad x \notin \left\{-\frac{1}{7}\right\}$$

The variables found in your answer were: $[x]$

Notes:

- Students' answers have mathematical content
- Current rudimentary interface for line by line working.
- (Started with CAS, not ATP!)

School exams

(Nadine Köcher & Chris Sangwin, 2014)

International Baccalaureate examinations in STACK?

	# marks	
(i) Awarded by STACK (2014) *exactly*	112	18%
(ii) Final answers and implied method marks	227	37%
(iii) Reasoning by equivalence	218	36%
Total of max of (ii) and (iii) per question	376	61%

School exams

(Nadine Köcher & Chris Sangwin, 2014)

International Baccalaureate examinations in STACK?

	# marks	
(i) Awarded by STACK (2014) *exactly*	112	18%
(ii) Final answers and implied method marks	227	37%
(iii) Reasoning by equivalence	218	36%
Total of max of (ii) and (iii) per question	376	61%

Repeat analysis with SQA Higher 2015.

	# marks	
(i) Awarded by STACK (v4.2) *exactly*	47	36%
(ii) Of which reasoning by equivalence	35	27%

Reasoning by equivalence

Work line by line: adjacent lines are "equivalent".

Reasoning by equivalence

Work line by line: adjacent lines are “equivalent".

$$
\begin{aligned}
&\log_3(x+17)-2 &&= \log_3(2x) \quad (x>0, x>-17)\\
\Leftrightarrow\ &\log_3(x+17)-\log_3(2x) &&= 2\\
\Leftrightarrow\ &\log_3\left(\frac{x+17}{2x}\right) &&= 2\\
\Leftrightarrow\ &\frac{x+17}{2x} &&= 3^2 = 9\\
\Leftrightarrow\ &x+17 &&= 18x\\
\Leftrightarrow\ &x &&= 1.
\end{aligned}
$$

Reasoning by equivalence

Work line by line: adjacent lines are “equivalent".

$$
\begin{aligned}
&\log_3(x+17)-2 &&= \log_3(2x) \quad (x>0, x>-17)\\
\Leftrightarrow &\log_3(x+17)-\log_3(2x) &&= 2\\
\Leftrightarrow &\log_3\left(\frac{x+17}{2x}\right) &&= 2\\
\Leftrightarrow &\frac{x+17}{2x} &&= 3^2 = 9\\
\Leftrightarrow &x+17 &&= 18x\\
\Leftrightarrow &x &&= 1.
\end{aligned}
$$

The above is a *single mathematical entity: the argument.*

Reasoning by equivalence

Work line by line: adjacent lines are "equivalent".

$$
\begin{aligned}
&\log_3(x+17)-2 &&= \log_3(2x) \quad (x>0, x>-17)\\
\Leftrightarrow\, &\log_3(x+17)-\log_3(2x) &&= 2\\
\Leftrightarrow\, &\log_3\left(\frac{x+17}{2x}\right) &&= 2\\
\Leftrightarrow\, &\frac{x+17}{2x} &&= 3^2 = 9\\
\Leftrightarrow\, &x+17 &&= 18x\\
\Leftrightarrow\, &x &&= 1.
\end{aligned}
$$

The above is a *single mathematical entity: the argument.*
The above is a single (long) *English sentence.*

What mathematical moves are included?

For this study

1. Algebraic equivalence of expressions

$$p \equiv q \Leftrightarrow p(x) = q(x), \quad \forall x \in X.$$

2. Equivalence of equations
   Same solutions: $V(p) = \{x \in X | p(x) = 0\}$.

What mathematical moves are included?

For this study

1. Algebraic equivalence of expressions
$$p \equiv q \Leftrightarrow p(x) = q(x), \quad \forall x \in X.$$
2. Equivalence of equations
Same solutions: $V(p) = \{x \in X | p(x) = 0\}$.
3. Equating coefficients.

What mathematical moves are included?

For this study

1. Algebraic equivalence of expressions

$$p \equiv q \Leftrightarrow p(x) = q(x), \quad \forall x \in X.$$

2. Equivalence of equations
   Same solutions: $V(p) = \{x \in X | p(x) = 0\}$.
3. Equating coefficients.
4. Support for Boolean connectives

$$(x-2)(x-3) = 0$$

$$x = 2 \text{ or } x = 3$$

What mathematical moves are included?

For this study

1. Algebraic equivalence of expressions

$$p \equiv q \Leftrightarrow p(x) = q(x), \quad \forall x \in X.$$

2. Equivalence of equations
   Same solutions: $V(p) = \{x \in X | p(x) = 0\}$.
3. Equating coefficients.
4. Support for Boolean connectives

$$(x-2)(x-3) = 0$$

$$x = 2 \text{ or } x = 3$$

5. Simple systems of inequalities, and simultaneous inequalities.

What mathematical moves are included?

For this study

1. Algebraic equivalence of expressions

$$p \equiv q \Leftrightarrow p(x) = q(x), \quad \forall x \in X.$$

2. Equivalence of equations
   Same solutions: $V(p) = \{x \in X | p(x) = 0\}$.
3. Equating coefficients.
4. Support for Boolean connectives

$$(x-2)(x-3) = 0$$

$$x = 2 \text{ or } x = 3$$

5. Simple systems of inequalities, and simultaneous inequalities.
6. Automatic detection of calculus operations.
7. Evaluation of the previous line with "let $x = \ldots$"

Equation to expression switch

Is this an equation "to solve", or a chain of equivalent expressions?

$$\frac{1}{x^2+1} = \frac{1}{(x+\mathrm{i})\,(x-\mathrm{i})}$$

Equation to expression switch

Is this an equation “to solve”, or a chain of equivalent expressions?

$$\frac{1}{x^2+1} = \frac{1}{(x+\mathrm{i})\,(x-\mathrm{i})}$$

$$= \frac{1}{2\,\mathrm{i}}\left(\frac{1}{x-\mathrm{i}} - \frac{1}{x+\mathrm{i}}\right)$$

Q15a: Auto-detection of calculus

Use integration by parts to find $\int x\sin(3x)dx$.

```
int(x*sin(3*x),x)
= (-x)/3*cos(3*x)-int((-1)/3*cos(3*x),x)
= (sin(3*x)-3*x*cos(3*x))/9+c
```

$$\begin{array}{lrl} & & \int x\ \sin(3\,x)\ \mathrm{d}x \\ \checkmark & & = \frac{-x}{3}\cos(3\,x) - \int \frac{-1}{3}\cos(3\,x)\ \mathrm{d}x \\ \int \ldots \mathrm{d}x & & = \frac{\sin(3\,x)-3\,x\ \cos(3\,x)}{9} + c \end{array}$$

Importance of RE in mathematics education

Reasoning by equivalence is important for the following reasons.

1. Natural progression from number and algebra.

Importance of RE in mathematics education

Reasoning by equivalence is important for the following reasons.

1. Natural progression from number and algebra.
2. Start of proof & rigour
   (deductive geometry?)

Importance of RE in mathematics education

Reasoning by equivalence is important for the following reasons.

1. Natural progression from number and algebra.
2. Start of proof & rigour
(deductive geometry?)
3. Multi-step extended calculation.

Importance of RE in mathematics education

Reasoning by equivalence is important for the following reasons.

1. Natural progression from number and algebra.
2. Start of proof & rigour
   (deductive geometry?)
3. Multi-step extended calculation.
4. Contains logical reasoning.

Importance of RE in mathematics education

Reasoning by equivalence is important for the following reasons.

1. Natural progression from number and algebra.
2. Start of proof & rigour
   (deductive geometry?)
3. Multi-step extended calculation.
4. Contains logical reasoning.
5. Included in many methods, e.g. solving ODEs.

Importance of RE in mathematics education

Reasoning by equivalence is important for the following reasons.

1. Natural progression from number and algebra.
2. Start of proof & rigour
   (deductive geometry?)
3. Multi-step extended calculation.
4. Contains logical reasoning.
5. Included in many methods, e.g. solving ODEs.
6. Key part of many pure mathematics proofs
   - Induction step
   - ϵ-δ proofs.

Importance of RE in school mathematics

The most important single form of reasoning in school mathematics is reasoning by equivalence.

(1/3 of marks in the IB exams are awarded for RE.)

STACK interface V0.1

Let students work line by line without explicit warrants.

- .... because that is what they do on paper ....

STACK interface V0.1

Let students work line by line without explicit warrants.

- .... because that is what they do on paper ....
- .... and we let them.

Research question

To what extent can we implement a typical school mathematics examination paper using STACK?

Research question

To what extent can we implement a typical school mathematics examination paper using STACK?

2018 SQA Advanced Higher Mathematics examinations.

Research question

To what extent can we implement a typical school mathematics examination paper using STACK?

2018 SQA Advanced Higher Mathematics examinations.

- Taken annually by about 3500 students, or 6% of the cohort.
- Single three hour paper, worth 100 marks.
- Calculators are permitted.
- Students are required to answer all questions.

Materials: `https://www.sqa.org.uk`.

Research question

1. To what extent can the questions be implemented *exactly* using the STACK?

Research question

1. To what extent can the questions be implemented *exactly* using the STACK?
2. To what extent is reasoning by equivalence included?

Research question

1. To what extent can the questions be implemented *exactly* using the STACK?
2. To what extent is reasoning by equivalence included?
3. What other forms of reasoning/processes are used and can this be automated?

Research question

1. To what extent can the questions be implemented *exactly* using the STACK?
2. To what extent is reasoning by equivalence included?
3. What other forms of reasoning/processes are used and can this be automated?
4. What cannot be automated, now and possibly in any system in the foreseeable future?

Caveat

No attempt to design an alternative question which measures the same competence.

Results

	# marks	
(i) Awarded by STACK (v4.3)	61	61%
(ii) Of which reasoning by equivalence	31	31%
(iii) Calculus moves	6	
Which contribute to	(15)	15%

Proof questions

9. Prove directly that:
(a) the sum of any three consecutive integers is divisible by 3;
(b) any odd integer can be expressed as the sum of two consecutive integers.

Proof questions

9. Prove directly that:
(a) the sum of any three consecutive integers is divisible by 3;
(b) any odd integer can be expressed as the sum of two consecutive integers.

11(d) Explain why matrix P is not associated with rotation about the origin.

Proof questions

9. Prove directly that:
(a) the sum of any three consecutive integers is divisible by 3;
(b) any odd integer can be expressed as the sum of two consecutive integers.

11(d) Explain why matrix P is not associated with rotation about the origin.

Q13a Show [reasoning from a diagram] that ...

Other things we “could do”

Q9b Sketch the locus in the complex plane

Other things we “could do”

Q9b Sketch the locus in the complex plane

Q5 Special interface: Euclidean algorithm

Q16a Special interface: Gaussian elimination

Q14c: Simple let

(c) Let S represent the sum of the first n terms of this arithmetic sequence. Find the values of n for which $S = 144$.

```
S = (d*(n-1)*n)/2+a*n
Let d=-16
Let a=80
Let S=144
144 = 80*n-8*(n-1)*n
16*n^2-176*n+288 = 0
n = 2 or n = 9
```

$$
\begin{array}{rl}
& S = \frac{d\cdot(n-1)\cdot n}{2} + a\cdot n \\
& \text{Let } d = -16 \\
& \text{Let } a = 80 \\
& \text{Let } S = 144 \\
\Leftrightarrow & 144 = 80\cdot n - 8\cdot(n-1)\cdot n \\
\Leftrightarrow & 16\cdot n^2 - 176\cdot n + 288 = 0 \\
\Leftrightarrow & n = 2 \text{ or } n = 9
\end{array}
$$

The variables found in your answer were: $[S, a, d, n]$

Q2: Refer to previous lines....

Use partial fractions to find $\int \frac{3\,x-7}{x^2-2\,x-15} dx$.

Write $\frac{3\,x-7}{x^2-2\,x-15}$ in partial fraction form

```
(3*x-7)/(x^2-2*x-15) = (3*x-7)/((x-5)*(x+3))
(3*x-7)/((x-5)*(x+3)) = B/(x+3)+A/(x-5)
3*x-7 = (x-5)*B+(x+3)*A
3*x-7 = x*(B+A)-5*B+3*A
-7 = 3*A-5*B and 3 = B+A
-7 = 3*A-5*B and 15 = 5*B+5*A
-7 = 3*A-5*B and 8 = 8*A
A = 1 and B = 2
(3*x-7)/((x-5)*(x+3)) = 2/(x+3)+1/(x-5)
```

$$\frac{3\,x-7}{x^2-2\,x-15} = \frac{3\,x-7}{(x-5)\,(x+3)} \qquad x \notin \{-3, 5\}$$

$$? \quad \frac{3\,x-7}{(x-5)\,(x+3)} = \frac{B}{x+3} + \frac{A}{x-5}$$

$$\Leftrightarrow \quad 3\,x-7 = (x-5)\,B + (x+3)\,A$$

$$\Leftrightarrow \quad 3\,x-7 = x\,(B+A) - 5\,B + 3\,A$$

$$\equiv (\cdots ?x) \quad \begin{cases} -7 = 3\,A - 5\,B \\ 3 = B + A \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} -7 = 3\,A - 5\,B \\ 15 = 5\,B + 5\,A \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} -7 = 3\,A - 5\,B \\ 8 = 8\,A \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} A = 1 \\ B = 2 \end{cases}$$

$$? \quad \frac{3\,x-7}{(x-5)\,(x+3)} = \frac{2}{x+3} + \frac{1}{x-5} \qquad x \notin \{-3, 5\}$$

The variables found in your answer were: $[x, A, B]$

Q2: Refer to previous lines....

Use partial fractions to find $\int \frac{3\cdot x-7}{x^2-2\cdot x-15} dx$.

Write $\frac{3\cdot x-7}{x^2-2\cdot x-15}$ in partial fraction form

```
(3*x)/(x^2-2*x-15)-7/(x^2-2*x-15) = (3*x-7)/((x-5)*(x+3))
(3*x-7)/((x-5)*(x+3)) = B/(x+3)+A/(x-5)
3*x-7 = A*(x+3)+B*(x-5)
Let x=5
A=1
Let x=-3
B=2
(3*x-7)/((x-5)*(x+3)) = 2/(x+3)+1/(x-5)
```

$$\frac{3\cdot x}{x^2-2\cdot x-15} - \frac{7}{x^2-2\cdot x-15} = \frac{3\cdot x-7}{(x-5)\cdot(x+3)} \qquad x \notin \{-3, 5\}$$

$$? \quad \frac{3\cdot x-7}{(x-5)\cdot(x+3)} = \frac{B}{x+3} + \frac{A}{x-5}$$

$$\Leftrightarrow \quad 3\cdot x - 7 = A\cdot(x+3) + B\cdot(x-5)$$

$$\text{Let } x = 5$$

$$\Leftrightarrow \quad A = 1$$

$$\text{Let } x = -3$$

$$? \quad B = 2$$

$$? \quad \frac{3\cdot x-7}{(x-5)\cdot(x+3)} = \frac{2}{x+3} + \frac{1}{x-5} \qquad x \notin \{-3, 5\}$$

The variables found in your answer were: $[A, B, x]$

Q2: Refer to previous lines....

Use partial fractions to find $\int \frac{3\cdot x-7}{x^2-2\cdot x-15} dx$.

Write $\frac{3\cdot x-7}{x^2-2\cdot x-15}$ in partial fraction form

```
(3*x)/(x^2-2*x-15)-7/(x^2-2*x-15) = (3*x-7)/((x-5)*(x+3))
(3*x-7)/((x-5)*(x+3)) = B/(x+3)+A/(x-5)
3*x-7 = A*(x+3)+B*(x-5)
Let x=5
A=1
Let x=-3
B=2
(3*x-7)/((x-5)*(x+3)) = 2/(x+3)+1/(x-5)
```

$$\frac{3\cdot x}{x^2-2\cdot x-15} - \frac{7}{x^2-2\cdot x-15} = \frac{3\cdot x-7}{(x-5)\cdot(x+3)} \qquad x \notin \{-3, 5\}$$

$$? \quad \frac{3\cdot x-7}{(x-5)\cdot(x+3)} = \frac{B}{x+3} + \frac{A}{x-5}$$

$$\Leftrightarrow \quad 3\cdot x - 7 = A\cdot(x+3) + B\cdot(x-5)$$

$$\text{Let } x = 5$$

$$\Leftrightarrow \quad A = 1$$

$$\text{Let } x = -3$$

$$? \quad B = 2$$

$$? \quad \frac{3\cdot x-7}{(x-5)\cdot(x+3)} = \frac{2}{x+3} + \frac{1}{x-5} \qquad x \notin \{-3, 5\}$$

The variables found in your answer were: $[A, B, x]$

Slight reformulation to separate partial fractions from integration $\rightarrow$ nested sub-arguments.

Nature of the subject

Polya 1962: *Mathematical Discovery: on understanding, learning and teaching problem solving.*

Nature of the subject

Polya 1962: *Mathematical Discovery: on understanding, learning and teaching problem solving.*

Patterns of thought for solving problems

- the pattern of two loci

Nature of the subject

Polya 1962: *Mathematical Discovery: on understanding, learning and teaching problem solving.*

Patterns of thought for solving problems

- the pattern of two loci
- superposition

Nature of the subject

Polya 1962: *Mathematical Discovery: on understanding, learning and teaching problem solving.*

Patterns of thought for solving problems

- the pattern of two loci
- superposition
- recursion

Nature of the subject

Polya 1962: *Mathematical Discovery: on understanding, learning and teaching problem solving.*

Patterns of thought for solving problems

- the pattern of two loci
- superposition
- recursion
- Cartesian pattern

Nature of the subject

Polya 1962: *Mathematical Discovery: on understanding, learning and teaching problem solving.*

Patterns of thought for solving problems

- the pattern of two loci
- superposition
- recursion
- Cartesian pattern

Legitimate patterns of thought $\rightarrow$ an acceptable proof.

Cartesian pattern

Descartes' *Rules for the Direction of the mind*.

1. Reduce any kind of problem to a mathematical problem.

Cartesian pattern

Descartes' *Rules for the Direction of the mind*.

1. Reduce any kind of problem to a mathematical problem.
2. Reduce any mathematical problem to algebra.

Cartesian pattern

Descartes' *Rules for the Direction of the mind*.

1. Reduce any kind of problem to a mathematical problem.
2. Reduce any mathematical problem to algebra.
3. Reduce any algebra problem to a single equation & solve.

Cartesian pattern

Descartes' *Rules for the Direction of the mind*.

1. Reduce any kind of problem to a mathematical problem.
2. Reduce any mathematical problem to algebra.
3. Reduce any algebra problem to a single equation & solve.

Polya: *"The more you know, the more gaps you can see in this project"*

Recursion

Polya's maxim:

if you cannot solve a problem, then solve a simpler one!

Find an explicit formula for S_n.
E.g. $S_n = 1 + 3 + 9 + 27 + \cdots + 3^{n-1}$.

Recursion

Polya's maxim:

if you cannot solve a problem, then solve a simpler one!

Find an explicit formula for S_n.
E.g. $S_n = 1 + 3 + 9 + 27 + \cdots + 3^{n-1}$.
Find an explicit formula for $S_n - S_{n-1}$.

Recursion

Polya's maxim:

if you cannot solve a problem, then solve a simpler one!

Find an explicit formula for S_n.
E.g. $S_n = 1 + 3 + 9 + 27 + \cdots + 3^{n-1}$.
Find an explicit formula for $S_n - S_{n-1}$.

Generalize from patterns $\rightarrow$ formal proof by induction.

De Morgan 1838

Example 1.—The sum of any number of successive odd numbers, beginning from unity, is a square number, namely, the square of half the even number which follows the last odd number. Let this proposition be true in any one single instance; that is, n being some whole number, let 1, 3, 5, ... up to $2n+1$ put together give $(n+1)^2$. Then the next odd number being $2n+3$, the sum of all the odd numbers up to $2n+3$ will be $(n+1)^2+2n+3$, or n^2+4n+4, or $(n+2)^2$. But $n+2$ is the half of the even number next following $2n+3$: consequently, if the proposition be true of any one set of odd numbers, it is true of one more. But it is true of the first odd number 1, for this is the square of half the even number next following. Consequently, being true of 1, it is true of $1+3$; being true of $1+3$, it is true of $1+3+5$; being true of $1+3+5$, it is true of $1+3+5+7$, and so on, *ad infinitum.*

De Morgan (1836)

Q12: Assessment of induction in STACK?

Prove by induction that, for all positive integers n,

$$\sum_{r=1}^{n} 3^{r-1} = \frac{3^n - 1}{2}.$$

```
"Let p(n) be the statement"
sum(3^(r-1),r,1,n) = (3^n-1)/2
Let n=1
sum(3^(r-1),r,1,1) = (3^1-1)/2
true
"Consider"
sum(3^(r-1),r,1,n+1)
= sum(3^(r-1),r,1,n)+3^n
= (3^n-1)/2+3^n
= (3*3^n-1)/2
= (3^(n+1)-1)/2
"and so"
sum(3^(r-1),r,1,n+1) = (3^(n+1)-1)/2
"which proves p(n) => p(n+1)."
```

Let p(n) be the statement

$\sum_{r=1}^{n} 3^{r-1} = \frac{3^n-1}{2}$

Let $n = 1$

$\Leftrightarrow$ $\sum_{r=1}^{1} 3^{r-1} = \frac{3^1-1}{2}$

$\Leftrightarrow$ **true**

Consider

$\sum_{r=1}^{n+1} 3^{r-1}$

✓ $= \sum_{r=1}^{n} 3^{r-1} + 3^n$

✓ $= \frac{3^n-1}{2} + 3^n$

✓ $= \frac{3 \cdot 3^n-1}{2}$

✓ $= \frac{3^{n+1}-1}{2}$

and so

$\sum_{r=1}^{n+1} 3^{r-1} = \frac{3^{n+1}-1}{2}$

which proves p(n) => p(n+1).

But what do students learn?

1. A love of intriguing patterns and tools for justifying them?

But what do students learn?

1. A love of intriguing patterns and tools for justifying them?
2. An incantation?

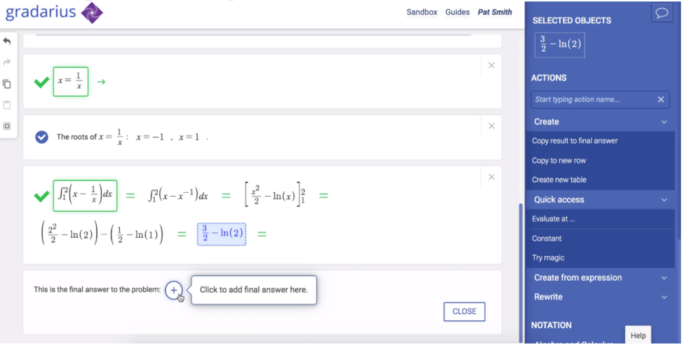
Free-text to pallet based input

Separation of assessment of

1. Legitimate forms of argument
2. correctness of algebraic steps within the argument?

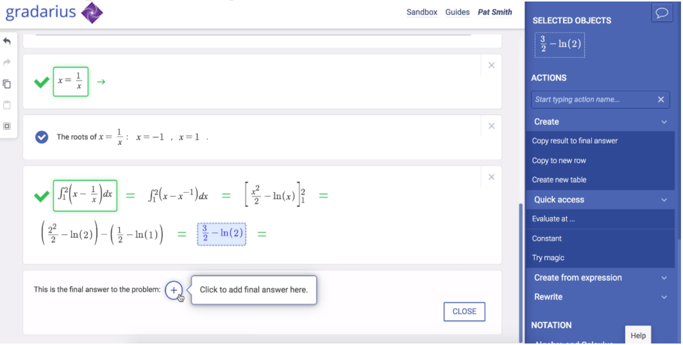
Many other learning systems

Gradarius: `http://www.gradarius.com/`



Many other learning systems

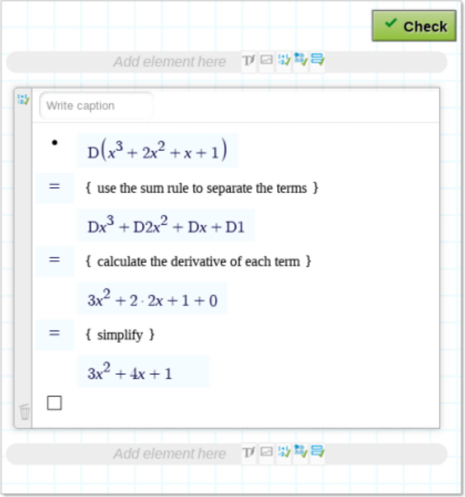
Gradarius: `http://www.gradarius.com/`



Also replicates current practice.

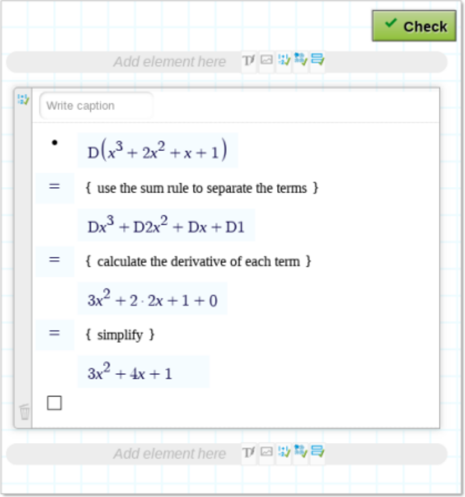
Structured derivations

Fourferries: `https://fourferries.com/`



Structured derivations

Fourferries: `https://fourferries.com/`



Structured derivation borrows from CS: more formality needed.

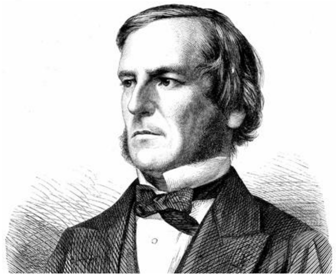
Reasoning → calculation has a long history

A “universal scientific language" would enable us to

> *judge immediately whether propositions presented to us are proved ... with the guidance of symbols alone, by a sure truly analytical method.*

Boole *Laws of thought* 1854

"to go under, over, and beyond" Aristotle's logic.



Resistance to change

To most people mathematics $=$ Stewart's *Calculus*.

Resistance to change

To most people mathematics $=$ Stewart's *Calculus*.
UK School textbooks....

Resistance to change

To most people mathematics $=$ Stewart's *Calculus*.
UK School textbooks....

> *The student is recommended to have as little as possible to do with* imaginary *quantities, that is, with quantities which have no meaning either as to* number *or* magnitude*. He need not wonder that the difficulties are likely to be introduced by the use of them, when he considers that* $\sqrt{-1}$ *signified an operation to be performed which is absolutely impossible. Any discussion upon the interpretation which* may *be give to such symbols, and the uses to which they may be applied, would be quite out of place in an Elementary Treatise like the present.*

Resistance to change

To most people mathematics $=$ Stewart's *Calculus*.
UK School textbooks....

> *The student is recommended to have as little as possible to do with* imaginary *quantities, that is, with quantities which have no meaning either as to* number *or* magnitude*. He need not wonder that the difficulties are likely to be introduced by the use of them, when he considers that* $\sqrt{-1}$ *signified an operation to be performed which is absolutely impossible. Any discussion upon the interpretation which* may *be give to such symbols, and the uses to which they may be applied, would be quite out of place in an Elementary Treatise like the present.*
> *Ed.* Lund 1841 *p. 75.*

Modified rules

(1) Multiplication does not retain equivalence.

$$CA = CB \Leftrightarrow A = B \vee C = 0. \tag{1}$$

$$CA = CB \wedge C \neq 0 \Leftrightarrow A = B \wedge C \neq 0. \tag{2}$$

$$A = B \Leftrightarrow (CA = CB \wedge C \neq 0) \vee A = B = 0. \tag{3}$$

Modified rules

(1) Multiplication does not retain equivalence.

$$CA = CB \Leftrightarrow A = B \vee C = 0. \tag{1}$$

$$CA = CB \wedge C \neq 0 \Leftrightarrow A = B \wedge C \neq 0. \tag{2}$$

$$A = B \Leftrightarrow (CA = CB \wedge C \neq 0) \vee A = B = 0. \tag{3}$$

(2) Powers and roots are evil.

Modified rules

(1) Multiplication does not retain equivalence.

$$CA = CB \Leftrightarrow A = B \vee C = 0. \tag{1}$$

$$CA = CB \wedge C \neq 0 \Leftrightarrow A = B \wedge C \neq 0. \tag{2}$$

$$A = B \Leftrightarrow (CA = CB \wedge C \neq 0) \vee A = B = 0. \tag{3}$$

(2) Powers and roots are evil.

$$\begin{aligned} A^2 = B^2 &\Leftrightarrow A^2 - B^2 = 0 \\ &\Leftrightarrow (A-B)(A+B) = 0 \\ &\Leftrightarrow A = B \vee A = -B. \end{aligned}$$

Modified rules

(1) Multiplication does not retain equivalence.

$$CA = CB \Leftrightarrow A = B \vee C = 0. \tag{1}$$

$$CA = CB \wedge C \neq 0 \Leftrightarrow A = B \wedge C \neq 0. \tag{2}$$

$$A = B \Leftrightarrow (CA = CB \wedge C \neq 0) \vee A = B = 0. \tag{3}$$

(2) Powers and roots are evil.

$$\begin{aligned} A^2 = B^2 &\Leftrightarrow A^2 - B^2 = 0 \\ &\Leftrightarrow (A - B)(A + B) = 0 \\ &\Leftrightarrow A = B \vee A = -B. \end{aligned}$$

(Auditing)
Fallacies in Mathematics, E. A. Maxwell (1959).

Student's comment

Sometimes STACK seems to have issues with answers that are essentially correct - (once I multiplied 2 square roots together i.e.`sqrt((x-3)*(x-5))` *and it said my answer was incorrect but then when I did* `sqrt(x-3)*sqrt(x-5)` *that was correct. It wasted time because I thought my calculation must have been wrong and was puzzled for a long time.*

Student's comment

Sometimes STACK seems to have issues with answers that are essentially correct - (once I multiplied 2 square roots together i.e.`sqrt((x-3)*(x-5))` *and it said my answer was incorrect but then when I did* `sqrt(x-3)*sqrt(x-5)` *that was correct. It wasted time because I thought my calculation must have been wrong and was puzzled for a long time.*

Is $\sqrt{ab} = \sqrt{a}\sqrt{b}$?

Resistance

Some people are very resistant to

1. Additional symbolism, e.g. “or“/$\vee$.

$$x = 2 \text{ or } x = 3$$

$$x = 2 \vee x = 3$$

Resistance

Some people are very resistant to

1. Additional symbolism, e.g. “or“/$\vee$.

$$x = 2 \text{ or } x = 3$$

$$x = 2 \vee x = 3$$

2. Change!

Resistance

Some people are very resistant to

1. Additional symbolism, e.g. “or“/$\vee$.

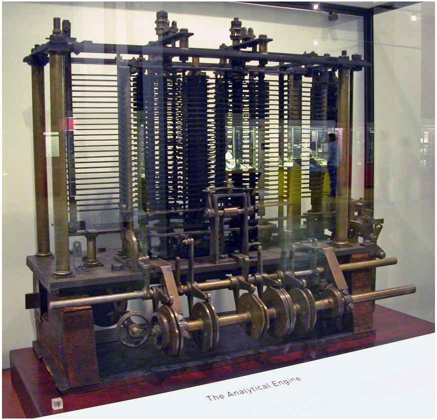
$$x = 2 \text{ or } x = 3$$

$$x = 2 \vee x = 3$$

2. Change!

Why not have a more formal layout for proofs?

Babbage and the Analytical Engine



Technology which looks back

Babbage set out to *print log tables!*

LOGARITHMIC SINES,

13 Deg.

′	Sine	Diff.	Cosec.	Tang.	Diff.	Cotang.	Secant	D.	Cosine	′
0	9·3520880	5469	10·6479120	9·3633641	5760	10·6366359	10·0112761	292	9·9887239	60
1	9·3526349	5461	10·6473651	9·3639401	5754	10·6360599	10·0113053	292	9·9886947	59
2	9·3531810	5454	10·6468190	9·3645155	5746	10·6354845	10·0113345	292	9·9886655	58
3	9·3537264	5446	10·6462736	9·3650901	5740	10·6349099	10·0113637	293	9·9886363	57
4	9·3542710	5440	10·6457290	9·3656641	5733	10·6343359	10·0113930	294	9·9886070	56
5	9·3548150	5432	10·6451850	9·3662374	5726	10·6337626	10·0114224	294	9·9885776	55
6	9·3553582	5425	10·6446418	9·3668100	5719	10·6331900	10·0114518	294	9·9885482	54
7	9·3559007	5419	10·6440993	9·3673819	5713	10·6326181	10·0114812	294	9·9885188	53
8	9·3564426	5410	10·6435574	9·3679532	5706	10·6320468	10·0115106	295	9·9884894	52
9	9·3569836	5404	10·6430164	9·3685238	5699	10·6314762	10·0115401	296	9·9884599	51
10	9·3575240	5397	10·6424760	9·3690937	5692	10·6309063	10·0115697	295	9·9884303	50
11	9·3580637	5390	10·6419363	9·3696629	5686	10·6303371	10·0115992	296	9·9884008	49
12	9·3586027	5382	10·6413973	9·3702315	5679	10·6297685	10·0116288	297	9·9883712	48
13	9·3591409	5376	10·6408591	9·3707994	5673	10·6292006	10·0116585	297	9·9883415	47
14	9·3596785	5369	10·6403215	9·3713667	5666	10·6286333	10·0116882	297	9·9883118	46
15	9·3602154	5361	10·6397846	9·3719333	5659	10·6280667	10·0117179	298	9·9882821	45
16	9·3607515	5355	10·6392485	9·3724992	5653	10·6275008	10·0117477	298	9·9882523	44
17	9·3612870	5347	10·6387130	9·3730645	5646	10·6269355	10·0117775	298	9·9882225	43
18	9·3618217	5341	10·6381783	9·3736291	5639	10·6263709	10·0118073	299	9·9881927	42
19	9·3623558	5334	10·6376442	9·3741930	5633	10·6258070	10·0118372	299	9·9881628	41
20	9·3628892	5327	10·6371108	9·3747563	5627	10·6252437	10·0118671	300	9·9881329	40

Technology which looks back

Babbage set out to *print log tables!*

13 Deg. LOGARITHMIC SINES,

′	Sine	Diff.	Cosec.	Tang.	Diff.	Cotang.	Secant	D.	Cosine	′
0	9·3520880	5469	10·6479120	9·3633641	5760	10·6366359	10·0112761	292	9·9887239	60
1	9·3526349	5461	10·6473651	9·3639401	5754	10·6360599	10·0113053	292	9·9886947	59
2	9·3531810	5454	10·6468190	9·3645155	5746	10·6354845	10·0113345	292	9·9886655	58
3	9·3537264	5446	10·6462736	9·3650901	5740	10·6349099	10·0113637	293	9·9886363	57
4	9·3542710	5440	10·6457290	9·3656641	5733	10·6343359	10·0113930	294	9·9886070	56
5	9·3548150	5432	10·6451850	9·3662374	5726	10·6337626	10·0114224	294	9·9885776	55
6	9·3553582	5425	10·6446418	9·3668100	5719	10·6331900	10·0114518	294	9·9885482	54
7	9·3559007	5419	10·6440993	9·3673819	5713	10·6326181	10·0114812	294	9·9885188	53
8	9·3564426	5410	10·6435574	9·3679532	5706	10·6320468	10·0115106	295	9·9884894	52
9	9·3569836	5404	10·6430164	9·3685238	5699	10·6314762	10·0115401	296	9·9884599	51
10	9·3575240	5397	10·6424760	9·3690937	5692	10·6309063	10·0115697	295	9·9884303	50
11	9·3580637	5390	10·6419363	9·3696629	5686	10·6303371	10·0115992	296	9·9884008	49
12	9·3586027	5382	10·6413973	9·3702315	5679	10·6297685	10·0116288	297	9·9883712	48
13	9·3591409	5376	10·6408591	9·3707994	5673	10·6292006	10·0116585	297	9·9883415	47
14	9·3596785	5369	10·6403215	9·3713667	5666	10·6286333	10·0116882	297	9·9883118	46
15	9·3602154	5361	10·6397846	9·3719333	5659	10·6280667	10·0117179	298	9·9882821	45
16	9·3607515	5355	10·6392485	9·3724992	5653	10·6275008	10·0117477	298	9·9882523	44
17	9·3612870	5347	10·6387130	9·3730645	5646	10·6269355	10·0117775	298	9·9882225	43
18	9·3618217	5341	10·6381783	9·3736291	5639	10·6263709	10·0118073	299	9·9881927	42
19	9·3623558	5334	10·6376442	9·3741930	5633	10·6258070	10·0118372	299	9·9881628	41
20	9·3628892	5327	10·6371108	9·3747563	5627	10·6252437	10·0118671	300	9·9881329	40

Knuth set out to *replicate movable type!*

STACK currently implements current practice (by design).

STACK currently implements current practice (by design).

In this sense it looks back.

Proof: Assessment of whole argument

Will require a sea-change in how we write mathematics.

Proof: Assessment of whole argument

Will require a sea-change in how we write mathematics.

"Those who cannot remember the past are condemned to repeat it."
(George Santayana)

Better interface

Firſt then, by D and E (that is, the Sum and Difference of two quantities) given, find F,G,T,R,(that is,the Product, Quotient, and Sum and Difference of the Squares.)

$a = ?$	1	$a+b=D$ } as in the Probleme.
$b = ?$	2	$a-b=E$
$1+2$	3	$2a=D+E$
$3 ,, \overline{2}$	4	$A=\frac{D+E}{2}=\mathcal{A}$.
$1-2$	5	$2b=D-E$
$5 \div \overline{2}$	6	$B=\frac{D-E}{2}=B$
or $1-4$	7	$B=\frac{D-E}{2}=B$

And ſo A and B are both explained by D & E,which was firſt to be found. The reſt follows eaſily.

Pell (1668)

Better interface

58 *Reſolution of Problemes.*

Firſt then, by D and E (that is, the Sum and Difference of two quantities) given, find F,G,T,R, (that is, the Product, Quotient, and Sum and Difference of the Squares.)

$a=?$	1	$a+b=D$ } as in the Probleme.
$b=?$	2	$a-b=E$
$1+2$	3	$2a=D+E$
$3 \,,, 2$	4	$A=\frac{D+E}{2}=A.$
$1-2$	5	$2b=D-E$
$5 \div 2$	6	$B=\frac{D-E}{2}=B$
or $1-4$	7	$B=\frac{D-E}{2}=B$

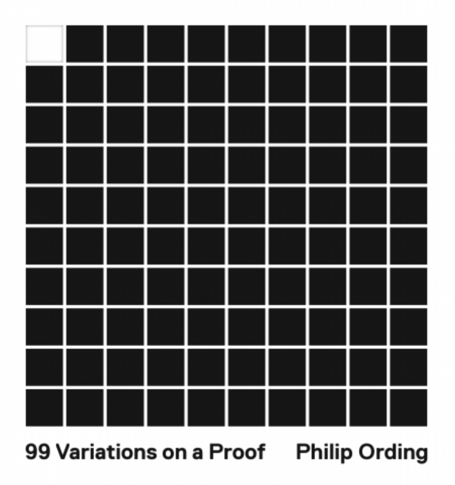
And ſo A and B are both explained by D & E, which was firſt to be found. The reſt follows eaſily.

Pell (1668)

(Ongoing work with M. Alafarj (2019)...)

Ording....

“... a deep and thoughtful examination of the nature of mathematical arguments, of mathematical style, and of proof itself.”



Conclusion

- We can automate assessment of a significant portion of school exams.

Conclusion

- We can automate assessment of a significant portion of school exams.
- If we “write to the format” we can do a lot more.

Conclusion

- We can automate assessment of a significant portion of school exams.
- If we “write to the format” we can do a lot more.
- To assess full proofs we need to re-engineer some aspects of teaching.

Conclusion

- We can automate assessment of a significant portion of school exams.
- If we "write to the format" we can do a lot more.
- To assess full proofs we need to re-engineer some aspects of teaching.
- Change is difficult: start early & be gentle.