Algorithm and abstraction in formal mathematics

Heather Macbeth

Fordham University

IPAM
17 February 2023

Mathematicians have strong opinions about which proofs are beautiful.

Mathematicians used to formal proofs have strong opinions about which formal proofs are beautiful.

What is this aesthetic? Is it the same as on paper, or does it have differences?

Example 1: Classification of wallpaper groups

Lemma

Let $2 \leq p \leq q \leq r$ be natural numbers, with

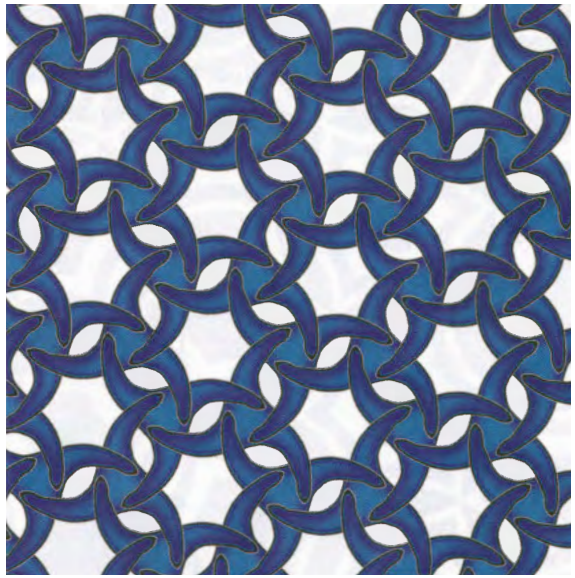
$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1. \qquad (\star)$$

Show that one of the following holds:

- $p = q = r = 3$
- $p = 2,\ q = r = 4$
- $p = 2,\ q = 3,\ r = 6$

$(3, 3, 3)$

$(2, 4, 4)$



$(2, 3, 6)$

An idiomatic paper proof

Proof.

We get $p = q = r = 3$ if all of $\frac{1}{p}, \frac{1}{q}$ and $\frac{1}{r}$ have their mean value of $\frac{1}{3}$.
Otherwise p must be 2.
If r and q have *their* mean value of $\frac{1}{4}$, we get $p = 2$, $q = r = 4$.
If not, q must be 3, and r is forced to be 6, by $(\star)$. □

(Adapted from Conway-Burgiel-Goodman-Strauss, *The Symmetries of Things*.)

An idiomatic formalized proof

Proof.

Fourier-Motzkin elimination on the inequalities $0 < \frac{1}{r} \leq \frac{1}{q} \leq \frac{1}{p} \leq \frac{1}{2}$ and the equality $(\star)$ yield that $\frac{1}{3} \leq \frac{1}{p}$ and $\frac{1}{4} \leq \frac{1}{q} < \frac{1}{2}$.

Since we are working with natural numbers, the inequality $\frac{1}{q} < \frac{1}{2}$ can be upgraded to $\frac{1}{q} \leq \frac{1}{3}$. Then Fourier-Motzkin elimination with this additional information yields that $\frac{1}{6} \leq \frac{1}{r}$.

We have now reduced to the finitely many cases $2 \leq p \leq 3$, $3 \leq q \leq 4$, $r \leq 6$, which can be checked against $(\star)$. □

(15 lines in Lean, written with Anne Baanen yesterday.)

Note: there is a similar lemma classifying triples $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} > 1$ which arises in the ADE classification of Dynkin diagrams, finite subgroups of $SO(3)$, ...

I chose to discuss the $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$ classification because it's epsilon harder.

Example 2: The Kochen-Specker paradox

Theorem

There does not exist a boolean function $s : \mathbb{R}^3 \to \{0, 1\}$*, such that for all triples* $u, v, w \in \mathbb{R}^3$ *of nonzero mutually-orthogonal vectors,* $s(u), s(v), s(w)$ *is* $0, 1, 1$ *in some order.*

Proof (Peres, 1991)[1]: Deduce a contradiction from the values of s on the following 33 nonzero vectors in $\mathbb{R}^3$:

$\bar{1}\bar{1}2$	$\bar{1}02$	$\bar{1}12$	$\bar{1}2\bar{1}$	$\bar{1}20$	$\bar{1}21$	$0\bar{1}2$	002	012	$02\bar{2}$	$02\bar{1}$
020	021	022	$1\bar{1}2$	102	112	$12\bar{1}$	120	121	$2\bar{2}0$	$2\bar{1}\bar{1}$
$2\bar{1}0$	$2\bar{1}1$	$20\bar{2}$	$20\bar{1}$	200	201	202	$21\bar{1}$	210	211	220

Here $\bar{1}$ is shorthand for -1, 2 is shorthand for $\sqrt{2}$, and $\bar{2}$ is shorthand for $-\sqrt{2}$. (Sorry!)

[1] This is an improvement on the original (1967) proof, which found a contradiction from a set of 117 vectors.

An idiomatic paper proof

The proof of the KS theorem entirely holds in table 1. In each line, the first ray, printed in boldface characters, is green. The second and third rays form, together with the first one, an orthogonal triad. Therefore they are red. Additional rays listed in the same line are also orthogonal to its first ray, therefore they too are red (only the rays that will be needed for further work are listed). When a red ray is printed in italic characters, this means that it is an 'old' ray, that was already found red in a preceding line. The choice of colours for the new rays appearing in each line is explained in the table itself.

Table 1. Proof of KS theorem in three dimensions.

Orthogonal triad			Other rays		The first ray is green because of
001	100	010	110	$1\bar{1}0$	choice of z axis
101	$\bar{1}01$	*010*			choice of x vs $-x$
011	$0\bar{1}1$	*100*			choice of y vs $-y$
$1\bar{1}2$	$\bar{1}12$	*110*	$\bar{2}01$	021	choice of x vs y
102	*$\bar{2}01$*	*010*	$\bar{2}11$		orthogonality to second and third rays
211	*$0\bar{1}1$*	*$\bar{2}11$*	$\bar{1}02$		orthogonality to second and third rays
201	*010*	*$\bar{1}02$*	$\bar{1}\bar{1}2$		orthogonality to second and third rays
112	*$1\bar{1}0$*	*$\bar{1}\bar{1}2$*	$0\bar{2}1$		orthogonality to second and third rays
012	*100*	*$0\bar{2}1$*	$1\bar{2}1$		orthogonality to second and third rays
121	*$\bar{1}01$*	*$1\bar{2}1$*	$0\bar{1}2$		orthogonality to second and third rays

The first, fourth and last lines contain rays 100, 021 and $0\bar{1}2$, respectively. These three rays are red and mutually orthogonal: this is the KS contradiction. It can be

(Peres, "Two simple proofs of the Kochen-Specker theorem", 1991.)

An idiomatic formalized proof

Proof (continued).

For each mutually-orthogonal triple u, v, w drawn from these 33 vectors, write down the statement

$$(\neg s(u) \wedge s(v) \wedge s(w)) \vee (s(u) \wedge \neg s(v) \wedge s(w)) \vee (s(u) \wedge s(v) \wedge \neg s(w)).$$

Case split as needed, and check that each of these 3^{large} cases gives a contradiction. □

(120 lines in HOL Light, written by John Harrison in 2005.[2])

[2] `https://github.com/jrh13/hol-light/blob/e736197400af93326c411dcc19d821f37a4f50cf/Tutorial/Custom_tactics.ml`

Example 3: Multiplication of Chebyshev polynomials

Let $T_n(x)$ denote the n-th Chebyshev polynomial of the first kind. Recall these polynomials satisfy a recurrence relation

$$T_{n+2}(x) = 2xT_{n+1}(x) - T_n(x).$$

Lemma

For all natural numbers m and k, $2T_mT_{m+k} = T_{2m+k} + T_k$.

Proof: ...

... **Inductive step** Suppose that the statement is true for m and $m+1$ (for all k). We will prove it for $m+2$.

An idiomatic human proof

Proof (continued).

Indeed,

$$\begin{aligned} 2T_{m+2}T_{m+k+2} &= 2[2xT_{m+1} - T_m]T_{m+k+2} \\ &= 2x[2T_{m+1}T_{(m+1)+(k+1)}] - 2T_mT_{m+(k+2)} \\ &= 2x[T_{2(m+1)+(k+1)} + T_{k+1}] - [T_{2m+(k+2)} + T_{k+2}] \\ &= [2xT_{2m+k+3} - T_{2m+k+2}] + [2xT_{k+1} - T_{k+2}] \\ &= T_{2m+k+4} + T_k. \end{aligned}$$

□

An idiomatic formalized proof [3]

Proof (continued).

Indeed, two applications of the inductive hypothesis give

$$2T_{m+1}T_{(m+1)+(k+1)} = T_{2(m+1)+(k+1)} + T_{k+1} \qquad (\star_1)$$

$$2T_mT_{m+(k+2)} = T_{2m+(k+2)} + T_{k+2} \qquad (\star_2)$$

and three applications of the recurrence relation give

$$T_{m+2} = 2xT_{m+1} - T_m \qquad (*_1)$$

$$T_{(2m+k+2)+2} = 2xT_{(2m+k+2)+1} - T_{2m+k+2} \qquad (*_2)$$

$$T_{k+2} = 2xT_{k+1} - T_k \qquad (*_3)$$

A Gröbner basis computation shows that LHS - RHS of the desired result,

$$2T_{m+2}T_{m+k+2} = T_{2m+k+4} + T_k,$$

is in the ideal generated by LHS - RHS of $(\star_1)$, $(\star_2)$, $(*_1)$, $(*_2)$, $(*_3)$. □

[3] https://hrmacbeth.github.io/computations_in_lean/02_Using_Polyrith.html#chebyshev-polynomials

Example 4: Lax-Milgram theorem

Let H be a real Hilbert space, $B : H \times H \to \mathbb{R}$ a bilinear form.

Theorem (Lax-Milgram)

Suppose there exist constants $\alpha, \beta > 0$ *so that*

- *(boundedness) for all* $u, v \in H$, $|B[u, v]| \leq \alpha \|u\| \|v\|$
- *(coercivity) for all* $u \in H$, $B[u, u] \geq \beta \|u\|^2$.

Then for each $f \in H^*$, *there exists a unique* $u \in H$ *so that for all* $v \in H$, $B[u, v] = f(v)$.

Proof: Construct a bounded linear map $A : H \to H$ such that for all $u, v \in H$, we have $B[u, v] = \langle A(u), v \rangle$. It suffices to show that A is bijective.

An idiomatic paper proof

3. **Claim 2.** A is injective, and the range of A, $R(A)$, is closed in H.

To see that A is injective, we use coercivity to write

$$\beta\|u\|^2 \leq B[u,u] = (Au,u) \overset{c.s.}{\leq} \|Au\|\|u\|.$$

If $\|u\| \neq 0$, we divide to see that

$$\|Au\| \geq \beta\|u\|.$$

(This is trivially true if $\|u\| = 0$.) In particular, if $u_1, u_2 \in H$, then $\|A(u_1 - u_2)\| \geq \beta\|u_1 - u_2\|$, from which it's clear that A is injective. (I.e., $u_1 \neq u_2 \implies Au_1 \neq Au_2$.)

To see that $R(A)$ is closed in H, let $\{u_j\}_{j=1}^{\infty} \subset H$ satisfy $Au_j \to w$ for some $w \in H$. We need to show that there exists $u \in H$ so that $Au = w$ (i.e., $w \in R(A)$).

For this, we notice that

$$\|u_i - u_j\| \leq \frac{1}{\beta}\|Au_i - Au_j\|.$$

The sequence $\{Au_j\}_{j=1}^{\infty}$ converges, so it must be Cauchy, so we see that $\{u_j\}_{j=1}^{\infty}$ must be Cauchy, and so must converge to some $u \in H$. Since A is bounded,

$$\|Au - w\| = \lim_{j\to\infty} \|Au - Au_j\| \leq \alpha \lim_{j\to\infty} \|u - u_j\| = 0.$$

I.e., $Au = w$. (Alternatively, since A is bounded, we know from the Closed Graph Theorem that A is closed, and this allows us to conclude $Au = w$.)

(Peter Howard, lecture notes at Texas A&M.)

An idiomatic formalized proof

Proof (continued).

By the coercivity of B, we have for all u

$$\beta\|u\|^2 \leq B[u,u] = \langle A(u), u\rangle \leq \|A(u)\|\|u\|,$$

so

$$\beta\|u\| \leq \|A(u)\|$$

(by the above if $u \neq 0$ and trivially if $u = 0$). Therefore A is antilipschitz, so it is injective and has closed range. □

(40 lines in Lean, written by Daniel Roca González in 2022.[4]

Theory of antilipschitz maps written by Yury Kudryashov in 2020.[5])

[4] `https://github.com/leanprover-community/mathlib/blob/master/src/analysis/inner_product_space/lax_milgram.lean`

[5] `https://leanprover-community.github.io/mathlib_docs/topology/metric_space/antilipschitz.html`

Example 5: smooth vector bundles

A smooth vector bundle with fibre F over a smooth manifold B consists of

- a disjoint union of topological vector spaces indexed by B
- a topology on the total space
- a collection of *trivializations*, each identifying the fibre-union over some open set $U \subseteq B$ homeomorphically with $U \times F$, commuting via projections with the identity on U, and fibrewise an isomorphism of topological vector spaces
- with the property that for two trivializations in the collection the induced map $U \cap V \to \mathrm{End}(F)$ is smooth.

Proposition

A smooth vector bundle is a smooth manifold.

An idiomatic paper proof

Proof.

Let H be the model space for the smooth manifold B. Given a trivialization $\psi = (\psi_b, \psi_f) : \pi^{-1}(U) \to U \times F$ and a chart $\varphi : V \xrightarrow{\sim} \varphi(V) \subseteq H$ for B, define a candidate chart

$$\Phi_{\psi,\varphi} : \pi^{-1}(U \cap V) \to \varphi(U \cap V) \times F$$

as

$$\Phi_{\psi,\varphi}(p) = (\varphi(\psi_b(p)), \psi_f(p)).$$

We need to check that for any two trivializations ψ_1, ψ_2 and any two charts φ_1, φ_2 the transition function $\Phi_{\psi_2,\varphi_2} \circ \Phi^{-1}_{\psi_1,\varphi_1}$ is smooth. This works out since $\psi_2 \circ \psi_1^{-1}$, φ_1 and φ_2 are all smooth. □

An idiomatic formalized proof

Proof.

Let H be the model space for the smooth manifold B. Let E be a smooth vector bundle over B with fibre F. We consider a sequence

$$E \dashrightarrow B \times F \dashrightarrow H \times F.$$

1. E is modelled on $B \times F$ with the charts being the trivializations, and with the transition functions between these charts lying in the "smooth fibrewise-linear" structure groupoid
2. $B \times F$ is in turn is modelled on $H \times F$ with the charts being the usual product manifold charts, and with the transition functions between these charts lying in the usual smooth manifold structure groupoid.

An idiomatic formalized proof (continued)

Proof (continued).

"Modellings" can be composed, so the modellings of E on $B \times F$ and of $B \times F$ on $H \times F$ yield a modelling of E on $H \times F$. Structure groupoid properties can also be composed, so the transition functions between these induced charts lie in the smooth manifold structure groupoid for $B \times F$. □

(Lean code, joint work with Floris van Doorn, not in mathlib yet.[6]

Theory of structure groupoids written by Sébastien Gouëzel in 2019.[7])

[6] `https://github.com/leanprover-community/mathlib/blob/d5a7ec38584eb449d96e1ff7c9d63bacca823d6e/src/geometry/manifold/vector_bundle/basic.lean`

[7] `https://leanprover-community.github.io/mathlib_docs/geometry/manifold/charted_space.html`

This is a workshop on machine-assisted *proof*. My examples in this talk have addressed the question of what constitutes a good proof of a fixed statement.

There is an orthogonal question: what the statements themselves should be, i.e. how to design the formalization of a whole theory. Some entry points to the literature here:

- "Competing inheritance paths in dependent type theory: a case study in functional analysis," Affeldt, Cohen, Kerjean, Mahboubi, Rouhling, Sakaguchi, 2020
- "A formalization of the change of variables formula for integrals in mathlib," Gouëzel, 2022