Without Loss of Generality

John Harrison

TPHOLs 2009, Munich

19th August 2009, 09:00–10:00

Without loss of generality

Mathematical proofs sometimes state that a certain assumption can be made ‘without loss of generality’ (WLOG).

Claims that proving the result in a more special case is nevertheless sufficient to justify the theorem in full generality.

Often justified by some sort of symmetry in the problem.

Example: Schur's inequality

Schur's asserts that for any nonnegative real numbers a, b and c and integer $k \geq 0$ one has

$$0 \leq a^k(a-b)(a-c) + b^k(b-a)(b-c) + c^k(c-a)(c-b)$$

A typical proof (e.g. the one on Wikipedia) would start:

> Without loss of generality, let $a \leq b \leq c$.

Justification for this step

Since $\leq$ is a total order, the three numbers must be ordered somehow, i.e. we must have (at least) one of $a \leq b \leq c$, $a \leq c \leq b$, $b \leq a \leq c$, $b \leq c \leq a$, $c \leq a \leq b$ or $c \leq b \leq a$. But the theorem is completely symmetric between a, b and c, so each of these cases is just a version of the other with a change of variables, and we may as well just consider one of them.

Two subtly different variants

- ‘The other cases are similar and are left to the reader’ — might just program the theorem prover to do multiple cases.
- ‘Without loss of generality . . . ’ — we want to appeal to a general logical principle

In a programmable theorem prover we can easily do the first here, but in general cases there may be a large or even infinite number of cases.

HOL Light proof of Schur's inequality

Conclusions from this exercise

- We embodied the ‘WLOG’ reasoning in a general theorem
- We deployed it and could follow the hand proof quite directly.

Conclusions from this exercise

- We embodied the ‘WLOG’ reasoning in a general theorem
- We deployed it and could follow the hand proof quite directly.
- However, establishing the required symmetry in the specific problem is the weak spot.
- It was easy here, but in more complicated situations it’s likely to be beyond the capacity of automation.

WLOG reasoning in geometry

Geometry is rich in WLOG principles, reflecting the importance of property-preserving transformations:

- Klein's "Erlanger Programm" emphasizes the role of transformations and invariance under classes of transformations
- Noether's work connects physical conservation principles to invariance of physical laws under transformations.

Invariance under transformations is often used to pick a more convenient or intuitive coordinate system in proofs.

First attempt at WLOG reasoning in geometry

Transforming quantifiers

We want to apply the same transformation to the *quantified* variables.

This can be justified based only on the surjectivity of the transformation:

```
|- !f. (!y. ?x. f x = y)
       ==> (!P. (!x. P x) <=> (!x. P (f x))) /\
           (!P. (?x. P x) <=> (?x. P (f x)))
```

Translation by a is clearly surjective because if $x = y - a$ then $a + x = y$.

Transforming higher-order quantifiers

We can actually justify similar transformations on quantifiers over *sets* of points, and set abstractions defining such sets.

```
|- !f. (!y. ?x. f x = y)
       ==> (!P. (!x. P x) <=> (!x. P (f x))) /\
           (!P. (?x. P x) <=> (?x. P (f x))) /\
           (!Q. (!s. Q s) <=> (!s. Q (IMAGE f s))) /\
           (!Q. (?s. Q s) <=> (?s. Q (IMAGE f s))) /\
           (!P. {x | P x} = IMAGE f {x | P (f x)})
```

This allows us to apply similar WLOG reasoning to properties involving sets.

Checking invariance of basic properties

We keep a reference variable `invariant_under_translation` of invariance theorems, which users are encouraged to add to for each new geometric concept:

```
|- !a x y. dist (a + x,a + y) = dist (x,y)

|- !a s. connected (IMAGE (\x. a + x) s) <=> connected
```

Some really show how to ‘pull’ translation through a concept:

```
|- !a x y. midpoint (a + x,a + y) = a + midpoint (x,y)

|- !a s. convex hull IMAGE (\x. a + x) s =
        IMAGE (\x. a + x) (convex hull s)
```

GEOM_ORIGIN_TAC

Our tactic `GEOM_ORIGIN_TAC` now automates everything:

- Applies basic invariance theorem
- Systematically transforms all other quantifiers
- Applies invariance theorems in a bottom-up sweep to prove invariance.

Usually the user doesn't have to intervene at all, unless the invariance becomes problematic.

GEOM_ORIGIN_TAC examples

The general situation

The most obvious general approach is the following:

```
|- (!x. ?f. transform f /\ nice (f x)) /\
   (!f x. transform f ==> (P (f x) <=> P x))
   ==> ((!x. P x) <=> (!x. nice x ==> P x))
```

The general situation

The most obvious general approach is the following:

```
|- (!x. ?f. transform f /\ nice (f x)) /\
   (!f x. transform f ==> (P (f x) <=> P x))
   ==> ((!x. P x) <=> (!x. nice x ==> P x))
```

We actually prefer the following slightly less obvious variant

```
|- (!x. ?f y. transform f /\ nice y /\ f y = x) /\
   (!f x. transform f /\ nice x ==> (P (f x) <=> P' x))
   ==> ((!x. P x) <=> (!y. nice y ==> P' y))
```

Invariance under linear transformations

A function $f : \mathbb{R}^M \to \mathbb{R}^N$ is linear iff:

```
|- linear f <=>
     (!x y. f (x + y) = f x + f y) /\ (!c x. f (c % x)
```

Many geometric properties satisfy preservation or pulling property under linear transformations:

```
|- !f a b. linear f ==> midpoint (f a,f b) = f (midpoin
```

```
|- !f s. linear f
          ==> convex hull IMAGE f s = IMAGE f (convex hu
```

Some require additional properties like injectivity:

```
|- !f s. linear f /\ (!x y. f x = f y ==> x = y)
         ==> (coplanar (IMAGE f s) <=> coplanar s)
```

Orthogonal transformations

Many properties are only preserved in general by *orthogonal transformations*, or more generally norm-preserving linear maps $f : \mathbb{R}^M \to \mathbb{R}^N$:

```
|- !f s. linear f /\ (!v. norm(f v) = norm v)
        ==> measure (IMAGE f s) = measure s

|- !f a b c.
       linear f /\ (!x. norm(f x) = norm x)
       ==> angle (f a,f b,f c) = angle (a,b,c)
```

Rotations

A couple even require a true *rotation*, i.e. an orthogonal transformation $\mathbb{R}^N \to \mathbb{R}^N$ whose matrix has determinant 1, namely cross products in $\mathbb{R}^3$:

```
|- !f x y. linear f /\
           (!x. norm(f x) = norm x) /\
           det(matrix f) = &1)
           ==> f x cross f y = f (x cross y)
```

and complex quotients in $\mathbb{R}^2$

```
|- !f w z. linear f /\
           (!x. norm(f x) = norm x) /\
           det(matrix f) = &1)
           ==> f w / f z = w / z
```

GEOM_BASIS_MULTIPLE_TAC

We have a similar list of invariance properties for linear transformations, `invariant_under_linear`.

It is used by `GEOM_BASIS_MULTIPLE_TAC`, which picks a rotation to bring a chosen vector onto the ‘positive’ part of any chosen coordinate axis.

```
|- !a k. 1 <= k /\ k <= dimindex(:N)
       ==> ?b f. orthogonal_transformation f /\
                 (2 <= dimindex(:N)
                  ==> det(matrix f) = &1) /\
                 f (b % basis k) = a /\
                 &0 <= b
```

We can’t in general pick a true rotation in 1 dimension, but the properties that require it are not 1-dimensional anyway.

An extended example

A ‘reflective’ version?

We never formalized the class of geometric properties that are invariant under various transformations. We can extend this class ‘ad hoc’ by adding new invariance theorems.

However, we could consider formalizing this class internally.

- More complicated and inefficient, may need to be revised repeatedly as new concepts are defined.
- May be able to deploy more interesting meta-properties that cannot be realized by the simple recursive transformations we use.

For example, a first-order assertion over vectors with M vector variables, even if the pattern of quantification is involved, can be reduced to spaces of dimension $\leq M$ [SAH].

Conclusions

- WLOG reasoning can be tricky to formalize.
- We can in general capture the key ideas in schematic theorems, but proving that the conditions are satisfied in specific cases may be difficult.
- Our systematic approach with automated support makes deploying such reasoning in geometry very straightforward.
- Interesting to consider many other ‘symmetry groups’.