A calculational proof of Andrew's challenge

David Gries[1]
Computer Science, Cornell University
Ithaca, NY 14853[2]

August 1996

At the Marktoberdorf summer school in August 1996, Larry Paulson lectured on his mechanical theorem prover, Isabelle, Natarajan Shankar lectured on his mechanical theorem prover, PVS, and I lectured on calculational logic. Both Paulson and Shankar suggested I try the calculational approach on Andrew's challenge, which is to prove theorem (1), given below, and after the summer school, Paulson emailed me F.J. Pelletier's collection of problems in first-order logic, which included Andrew's challenge.[3]

$$(1)\quad ((\exists x \forall y \mid : p.x \equiv p.y) \ \equiv \ ((\exists x \mid : q.x) \ \equiv \ (\forall y \mid : p.y))) \ \equiv$$
$$((\exists x \forall y \mid : q.x \equiv q.y) \ \equiv \ ((\exists x \mid : p.x) \ \equiv \ (\forall y \mid : q.y)))$$

In proving Andrew's challenge using the calculational approach, I use theorems given in the text [1] (or in its as-yet-unpublished second edition). The Appendix contains theorems used here that may be unfamiliar to the reader.

Now, $\equiv$ is both associative and symmetric, so we can rewrite Andrew's challenge as

$$P \ \equiv \ Q$$

where P and Q are defined by the following.

$$P: \ (\exists x \forall y \mid : p.x \equiv p.y) \ \equiv \ (\exists x \mid : p.x) \ \equiv \ (\forall y \mid : p.y)$$
$$Q: \ (\exists x \forall y \mid : q.x \equiv q.y) \ \equiv \ (\exists x \mid : q.x) \ \equiv \ (\forall y \mid : q.y)$$

This form gives us the impression that perhaps P is valid (or invalid), regardless of p. If this is the case, then Q is also valid (or invalid). Hence, we try to prove P.

---

[1]Supported by NSF grants CDA-9214957 and CCR-9503319.

[2]http://www.cs.cornell.edu/Info/People/gries/gries.html gries@cs.cornell.edu

[3]We use the notation $(\forall x \mid : P)$ instead of the more traditional $\forall x.P$; the reasons for this are explained in [1]. $(\forall x \mid : (\exists y \mid : P))$ may be abbreviated as $(\forall x \exists y \mid : P)$. Also, we use $\equiv$ for equality over the bools and $=$ for equality over any type (including the bools). Our precedences are, beginning with the tightest, $\neg$, $=$, $\vee$ and $\wedge$, $\Rightarrow$ and $\Leftarrow$, $\equiv$. Finally, in order to eliminate parentheses, we write $p.x$ instead of $p(x)$ for application of function p to variable x.

We don't have many theorems that deal with $\equiv$ as they appear in P, so we try to prove P by mutual implication, proving instead

(2) $((\exists x \forall y \mid: p.x \equiv p.y) \;\equiv\; (\exists x \mid: p.x)) \;\Leftarrow\; (\forall y \mid: p.y)$ and
(3) $((\exists x \forall y \mid: p.x \equiv p.y) \;\equiv\; (\exists x \mid: p.x)) \;\Rightarrow\; (\forall y \mid: p.y)$

We prove (2):

Assume $(\forall y \mid: p.y)$
$(\exists x \forall y \mid: p.x \equiv p.y) \;\equiv\; (\exists x \mid: p.x)$
$=$ $\langle$Assumption, instantiated with $y := x$ and with $y := y$,
so $p.x \;\equiv\; true$ and $p.y \;\equiv\; true \rangle$
$(\exists x \forall y \mid: true \equiv true) \;\equiv\; (\exists x \mid: true)$
$=$ $\langle$Identity of $\equiv$ (4); $(\forall y \mid: true) \;\equiv\; true \rangle$
$(\exists x \mid: true) \;\equiv\; (\exists x \mid: true)$ —Reflexivity of $\equiv$ (5)

We prove (3).

(3)
$=$ $\langle$Contrapositive, $X \Rightarrow Y \;\equiv\; \neg Y \Rightarrow \neg X \rangle$
$\neg(\forall y \mid: p.y) \;\Rightarrow\; \neg((\exists x \forall y \mid: p.x \equiv p.y) \;\equiv\; (\exists x \mid: p.x))$
$=$ $\langle$De Morgan (11) on antecedent;
$\neg(X \equiv Y) \;\equiv\; X \;\equiv\; \neg Y$ and De Morgan (10) on the consequent$\rangle$
$(\exists y \mid: \neg p.y) \;\Rightarrow\; ((\exists x \forall y \mid: p.x \equiv p.y) \;\equiv\; (\forall x \mid: \neg p.x))$

By Metatheorem Witness (12), the last formula is a theorem iff the following one is.

$$\neg p.\hat{y} \;\Rightarrow\; ((\exists x \forall y \mid: p.x \equiv p.y) \;\equiv\; (\forall x \mid: \neg p.x))$$

We calculate:

Assume $\neg p.\hat{y}$, so also $p.\hat{y} \;\equiv\; false$
$(\exists x \forall y \mid: p.x \equiv p.y)$
$=$ $\langle$Zero of $\vee$ (6) on range of $\forall y$ —we're heading to change $p.x$ to $p.\hat{y} \rangle$
$(\exists x \mid: (\forall y \mid true \vee (y \equiv \hat{y}) : p.x \equiv p.y))$
$\langle$Range split (9); One-point rule (8)$\rangle$
$(\exists x \mid: (\forall y \mid: p.x \equiv p.y) \wedge p.x \equiv p.\hat{y})$
$=$ $\langle$Substitution (7)$\rangle$
$(\exists x \mid: (\forall y \mid: p.\hat{y} \equiv p.y) \wedge p.x \equiv p.\hat{y})$
$=$ $\langle$One-point rule; Range split; Zero of $\vee$ —eliminate the conjunct $p.x \equiv p.\hat{y} \rangle$
$(\exists x \forall y \mid: p.\hat{y} \equiv p.y)$
$=$ $\langle$Assumption $p.\hat{y} \equiv false$; $false \;\equiv\; X \;\equiv\; \neg X \rangle$
$(\exists x \forall y \mid: \neg p.y)$
$=$ $\langle$Provided x doesn't occur free in X, $(\exists x \mid: X) \;\equiv\; X \rangle$
$(\forall y \mid: \neg p.y)$

References

[1] Gries, D., and F.B. Schneider. *A Logical Approach to Discrete Math.* Springer Verlag, NY, 1993.

Appendix. Some of the theorems used in the proof

(4) **Identity of $\equiv$:** $\mathit{true} \equiv Q \equiv Q$

(5) **Reflexivity of $\equiv$:** $P \equiv P$

(6) **Zero of $\vee$:** $P \vee \mathit{true} \equiv \mathit{true}$

(7) **Substitution:** $X{=}Y \wedge E^V_X \equiv X{=}Y \wedge E^V_Y$

(8) **One-point rule:** Provided x does not occur free in E,

$(\forall x \mid x = E : P) = P[x := E]$

(9) **Range split:** $(\forall x \mid R \vee S : P) = (\forall x \mid R : P) \wedge (\forall x \mid S : P)$

(10)**De Morgan:** $\neg(\exists x \mid R : P) \equiv (\forall x \mid R : \neg P)$

(11)**De Morgan:** $\neg(\forall x \mid R : P) \equiv (\exists x \mid R : \neg P)$

(12) **Metatheorem Witness.** Suppose $\hat{x}$ does not occur free in P, Q, and R. Then

$(\exists x \mid R : P) \Rightarrow Q$ is a theorem iff
$(R \wedge P)[x := \hat{x}] \Rightarrow Q$ is a theorem.