Designing a proof of unique factorization

Introduction

In this note, we design a proof of "unique factorization", that is:

(0) "Two bags of primes are equal if their products are equal."

The impetus for our exposition is EWD755, in which Edsger Dijkstra recasts a proof of unique factorization (by Courant and Robbins) in a more formal nomenclature. Though we consider the formal setting to be an improvement, we feel that the structure of the argument and the heuristics can be improved further. Hence this note.

* *
*

A proof by induction

We proceed by mathematical induction. To preserve the symmetry between the bags, we induct over their "combined size", that is, the size of their union.

In the base case, the combined size is 0, hence both bags equal the empty bag. Thus the bags are equal, which establishes (0).

In the induction step, we assume the combined size to be at least 1, so that one of the bags has an element. To capture this asymmetry succinctly, we introduce names: We call the element p, we call the bag containing it B, and we call the other bag C. Phrased in this terminology, our demonstrandum (0) becomes:

(1) $*.B = *.C \quad \Rightarrow \quad B = C$,

under assumption of the induction hypothesis —that (0) holds for bags of smaller combined size than that of B and C— and $p \in B$.

Noting the symmetry between B and C in (1), we choose to restore the lost symmetry by assuming $p \in C$, deferring the proof obligation until the next section. On account of $p \in B$ and $p \in C$ we may now write:

(2) $B = b \cup \{p\}$ and $C = c \cup \{p\}$,

for bags b and c. We then calculate the induction step as follows:

$$
\begin{array}{ll}
 & B = C \\
\equiv & \quad \{ \ (2) \ \} \\
 & b \cup \{p\} \ = \ c \cup \{p\}
\end{array}
$$

$\Leftarrow$ { Leibniz }

$b = c$

$\Leftarrow$ { induction, since the combined size of b and c is smaller than the combined size of B and C }

$*.b \;=\; *.c$

$\equiv$ { arithmetic, $p \neq 0$ }

$(*.b) * p \;=\; (*.c) * p$

$\equiv$ { arithmetic }

$*.(b \cup \{p\}) \;=\; *.(c \cup \{p\})$

$\equiv$ { (2) }

$*.B \;=\; *.C$.

This completes the proof by induction. In the next section, we address our deferred proof obligation $p \in C$.

* * *

Proving $p \in C$

In the previous section, the proof obligation for the induction step was:

$$p \in B \;\Rightarrow\; (*.B \;=\; *.C \;\Rightarrow\; B = C) \quad .$$

However, in order to meet this proof obligation, we assumed $p \in C$. To reconcile this assumption, by predicate calculus it suffices to prove:

$$p \in B \;\wedge\; *.B \;=\; *.C \quad \Rightarrow \quad p \in C \quad .$$

(We could use the induction hypothesis as well, but we will not need to do so.)

Aiming for a more symmetric proof shape, we group the syntactically similar predicates using shunting:

$$*.B \;=\; *.C \;\Rightarrow\; (p \in B \;\Rightarrow\; p \in C) \quad .$$

This shape suggests that we construct a weakening chain from $p \in B$ to $p \in C$, using $*.B = *.C$. The standard way to use an equality is to "substitute equals for equals" (Leibniz) , which suggests the following proof sketch:

$p \in B$

$\Rightarrow$ { Lemma —to be determined— }

... p ... $(*.B)$ ...

$\equiv$ { $*.B = *.C$ }

... p ... $(*.C)$ ...

$\Rightarrow$ { converse of Lemma }

$p \in C$.

This completes the proof of $p \in C$. In the next section, we address the design and proof of our Lemma.

* * *

Designing the Lemma

The shape of the Lemma is:

$$p \in X \quad \Rightarrow \quad \text{... } p \text{ ... } (*.X) \text{ ...} \quad , \quad \text{for } X \text{ a bag of primes} \quad .$$

We prove the Lemma by refining its consequent under assumption of its antecedent. In the following calculation, $\sqsubseteq$ denotes the "divides" relation:

p

$=$ { heading for $*.X$, we use the one-point rule to introduce $*$ }

$*.\{p\}$

$\sqsubseteq$ { using $\{p\} \subseteq X$ $(\equiv p \in X)$, arithmetic }

$*.X$.

Thus our Lemma is:

$$p \in X \quad \Rightarrow \quad p \sqsubseteq *.X \quad .$$

This completes the design and proof of the Lemma. In the next section, we address the proof of its converse.

* * *

Proof of the converse of the Lemma

The converse of the Lemma is $p \in X \Leftarrow p \sqsubseteq *.X$. We begin our calculation with the consequent $p \sqsubseteq *.X$, because the antecedent $p \in X$ does not offer any manipulative possibilities. However, the calculation is most easily designed by manipulating both expressions towards each other. So that the reader can follow this development, we have labelled the steps in the direction and order that they should be read, and annotated the steps with heuristic guidance. (The calculation may also be read straightforwardly from top to bottom, omitting the heuristics.)

The double angle brackets $\langle\!\langle \ldots \rangle\!\rangle$ denote quantification over a bag, meaning that the dummy should range over every element of the bag, including duplicates. Such notation can easily lead to error, since a boolean expression like $q \in X$ does not capture multiplicity. (For example, $q \in \{2,2\} \equiv q \in \{2\}$.) To avoid error, we shall simply refrain from manipulating expressions like $q \in X$ in this way. But the reader should note that when the quantified operator is idempotent, multiplicity becomes irrelevant, and hence all the standard quantification rules apply!

$p \sqsubseteq *.X$

$\downarrow\mathbf{0}$ $\equiv$ { property of $*$ }

[**Heuristic:** *The target of this calculation is* $p \in X$. *By expanding* $*.X$ *into* $\langle\!\langle *q : q \in X : q \rangle\!\rangle$, *we form part of the target. It remains to bridge the gap from* $\langle\!\langle \ldots q \in X \ldots \rangle\!\rangle$ *to* $p \in X$, *which suggests use of the one-point rule. Since it is not clear how to apply the one-point rule at this stage, we turn to the bottom of the calculation.*]

$p \sqsubseteq \langle\!\langle *q : q \in X : q \rangle\!\rangle$

$\downarrow\mathbf{3}$ $\Rightarrow$ { p prime, see Appendix 1 }

[**Heuristic:** *In order to form an existential quantification and to bring* p *inside it, we posit a distributive property whose proof we defer to Appendix 1. (The proof uses that* p *is prime.)*]

$\langle\!\langle \exists q : q \in X : p \sqsubseteq q \rangle\!\rangle$

$\updownarrow\mathbf{4}$ $\Rightarrow$ { p, q prime, see Appendix 0 }

[**Heuristic:** *For the final step we posit* $p \sqsubseteq q \Rightarrow p = q$. *In Appendix 0 we prove this for all prime* p *and* q .]

$\langle\!\langle \exists q : q \in X : p = q \rangle\!\rangle$

$\uparrow\mathbf{2}$ $\equiv$ { predicate calculus }

[**Heuristic:** *We appeal to predicate calculus to put* $q \in X$ *into the range.*]

$\langle\!\langle \exists q : p = q : q \in X \rangle\!\rangle$

$\uparrow\mathbf{1}$ $\equiv$ { one-point rule, since $\exists$ is idempotent }

[**Heuristic:** *Applying the one-point rule yields* $\langle\!\langle \ldots q : p = q : q \in X \rangle\!\rangle$. *Type considerations restrict the choice of quantifier to* $\forall$ *and* $\exists$. *Furthermore, since in Step 0,* $q \in X$ *appears in the range, the quantifier should be symmetric in range and term. Hence we choose* $\exists$.]

$p \in X$.

This completes the proof of the converse of the Lemma, which relied on the following properties:

(3) $p \sqsubseteq q \;\Rightarrow\; p = q$, for p, q prime

(4) $p \sqsubseteq \langle\langle *q :: q \rangle\rangle \;\Rightarrow\; \langle\langle \exists q :: p \sqsubseteq q \rangle\rangle$, for p prime .

Readers familiar with these properties may stop here. In the following Appendices, we refine them using:

$$x \sqsubseteq p \;\Rightarrow\; x = 1 \;\vee\; x = p \qquad \text{and} \qquad p \neq 1$$

as the only properties of prime p .

* *
*

Appendix 0: Proof of (3)

⟦ Context: p prime, q prime

$p \sqsubseteq q$
⇒ { q prime }
$p = 1 \;\vee\; p = q$
≡ { p prime, hence $p \neq 1$ }
$p = q$
⟧

* *
*

Appendix 1: Proof of (4)

We prove (4) by induction over the size of the range of quantification. In the base case, both sides equivale **false** . The induction step boils down to "the case for 2" :

$$p \sqsubseteq (r * s) \;\Rightarrow\; p \sqsubseteq r \;\vee\; p \sqsubseteq s \quad .$$

(See JAW97 for more on this phenomenon.) We prove this case as follows:

$$
\begin{array}{ll}
 & p \sqsubseteq (r * s) \quad \Rightarrow \quad p \sqsubseteq r \ \vee \ p \sqsubseteq s \\
\equiv & \quad \{ \text{ predicate calculus } \} \\
 & (p \sqsubseteq (r * s) \ \Rightarrow \ p \sqsubseteq r) \quad \vee \quad p \sqsubseteq s \\
\Leftarrow & \quad \{ \text{ Euclid's lemma } \} \\
 & p \,\mathbf{gcd}\, s = 1 \quad \vee \quad p \sqsubseteq s \\
\equiv & \quad \{ \text{ lattice theory } \} \\
 & p \,\mathbf{gcd}\, s = 1 \quad \vee \quad p \,\mathbf{gcd}\, s = p \\
\Leftarrow & \quad \{ \ p \ \text{ prime } \} \\
 & p \,\mathbf{gcd}\, s \ \sqsubseteq \ p \\
\equiv & \quad \{ \text{ lattice theory } \} \\
 & \mathbf{true} \qquad .
\end{array}
$$

Readers interested in the proof and design of Euclid's Lemma:

$$p \,\mathbf{gcd}\, s = 1 \quad \Rightarrow \quad (\ p \sqsubseteq (r * s) \ \Rightarrow \ p \sqsubseteq r \ )$$

are referred to JAW10b .

* * *

Acknowledgements

We are indebted to Diethard Michaelis for extensive comments on an earlier draft.

Original version: Santa Cruz, 2 January 2007

Revision: Santa Cruz, 15 March 2007

Final revision: Santa Cruz, 15 November 2007

Apurva Mehta
apurva@mathmeth.com

Jeremy Weissmann
jeremy@mathmeth.com