

The Specification of a Generic Zip

Roland Backhouse and Paul Hoogendijk
Generic Programming Summer School
Oxford, August 2002

(Lower Order) Naturality

$$\text{zip.F.G} : G \bullet F \leftarrow F \bullet G \text{ .}$$

A zip is a *proper* natural transformation.

A zip transforms one structure to another without loss or duplication of values.

(Higher Order) Naturality

$$\text{zip.F} : (\bullet F) \leftarrow (F \bullet) \text{ .}$$

Categorical Nat Trans (Revision)

A natural transformation is an arrow in the functor category. I.e.,

$$\eta : F \leftarrow G$$

means that the following diagram commutes (for all A, B and $f : A \leftarrow B$)

$$\begin{array}{ccc}
 F.A & \xleftarrow{\eta_A} & G.A \\
 \uparrow F.f & & \uparrow G.f \\
 F.B & \xleftarrow{\eta_B} & G.B
 \end{array}$$

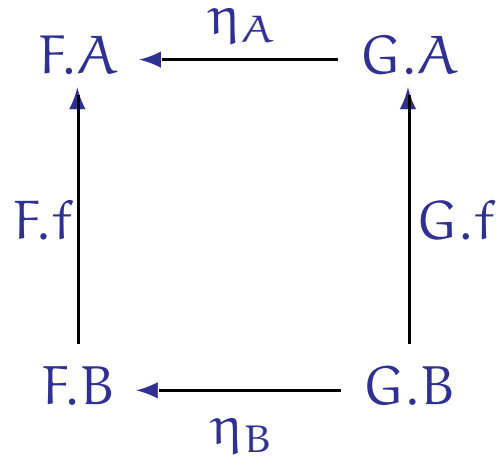
Now, if F is a functor, $(\bullet F)$ and $(F \bullet)$ are endofunctors on the functor category.

$(\bullet F)$ maps functor (object) G to $G \bullet F$ and natural transformation (arrow) η to $\eta \bullet F$, where $(\eta \bullet F)_A = \eta_{F.A}$.

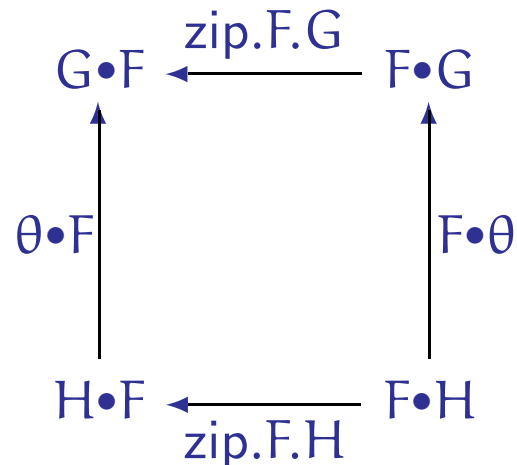
$(F \bullet)$ maps functor (object) G to $F \bullet G$ and natural transformation (arrow) η to $F \bullet \eta$, where $(F \bullet \eta)_A = F.(\eta_A)$.

Categorical NT Revision (Continued)

Diagram defining $\eta : F \leftarrow G$



instantiated for $\text{zip}.F : (\bullet F) \leftarrow (F \bullet)$



where $\theta : G \leftarrow H$ is a natural transformation.

Allegorical Naturality

Recall that parametricity was defined in terms of *relations*.

Recall also that, in the particular case that $\mathbf{t}$ has type $\langle \forall \alpha :: F.\alpha \leftarrow G.\alpha \rangle$, $\mathbf{t}$ is parametric is equivalent to $\mathbf{t}$ is a natural transformation (in the underlying category of maps).

This is a stroke of luck for functional programmers, BUT their luck has run out!

The equality in

$$(\theta \bullet F) \circ \text{zip}.F.H = \text{zip}.F.G \circ (F \bullet \theta)$$

is too severe — because

- θ may be nondeterministic.
- Zips are partial.

Nondeterminism

Take $F := \text{List}$ and $G = H := \times$.

$\text{zip}.F.H$ and $\text{zip}.F.G$ are both the inverse of conventional zips. They unzip a list of pairs to a pair of lists.

Take $\theta := \text{id} \cup \text{swap}$.

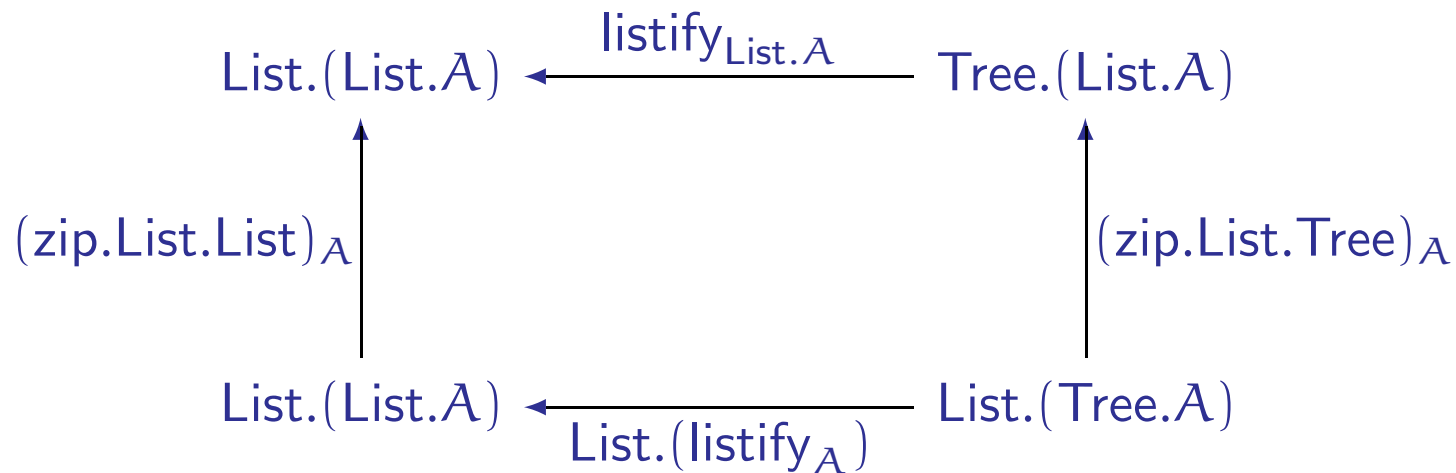
θ nondeterministically swaps the elements of a pair or not.

$(\theta \bullet F) \circ \text{zip}.F.H$ unzips a list of pairs into a pair of lists and swaps the lists or not.

$\text{zip}.F.G \circ (F \bullet \theta)$ first swaps some of the elements of a list of pairs and then unzips it into a pair of lists.

$$(\theta \bullet F) \circ \text{zip}.F.H \quad \subset \quad \text{zip}.F.G \circ (F \bullet \theta) \quad .$$

Partiality



View both paths through the diagram as partial relations of type $\text{List.}(\text{List.}A) \leftarrow \text{List.}(\text{Tree.}A)$.

The lower path (via $\text{List.}(\text{List.}A)$) includes the upper path (via $\text{Tree.}(\text{List.}A)$).

Reason: for the lower path, the sizes of the trees must be the same; for the upper path, the trees must have the same shape.

zip.F **is parametric.**

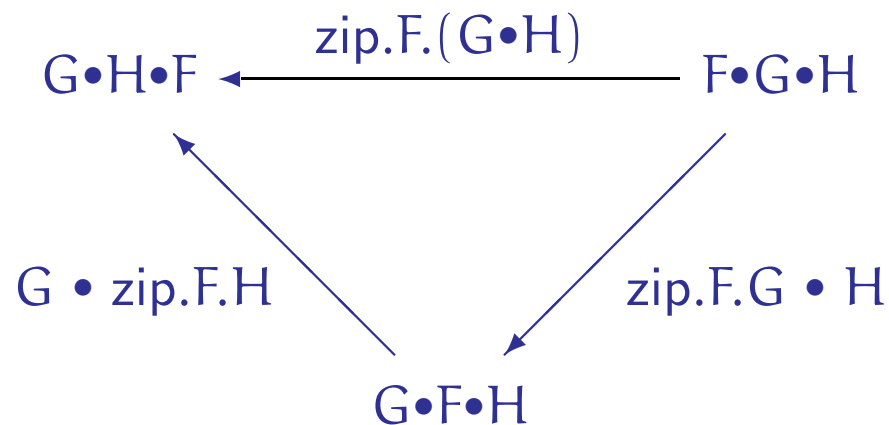
That is, for all $\theta : G \hookrightarrow H$,

$$(\theta \bullet F) \circ \text{zip.F.H} \subseteq \text{zip.F.G} \circ (F \bullet \theta) \text{ .}$$

Compositionality

Informally, zip.F is a monoid homomorphism.

(Note: more than this: zip.F should respect pointwise extension of relators. For full discussion see Hoogendijk's thesis.)



$$\text{zip.F.}(G \bullet H) = (G \bullet \text{zip.F.H}) \circ (\text{zip.F.G} \bullet H) \text{ .}$$

$$\text{zip.F.Id} = \text{id} \bullet F \text{ .}$$

Zips

Definition 1 (Half Zip) Consider a fixed relator F and a pointwise closed class of relators $\mathcal{G}$. Then the members of the collection $\text{zip}.F.G$, where G ranges over $\mathcal{G}$, are called *half-zips* iff

- (a) $\text{zip}.F.G : G \bullet F \leftarrow F \bullet G$, for each G in $\mathcal{G}$,
- (b) $(\theta \bullet F) \circ \text{zip}.F.H \subseteq \text{zip}.F.G \circ (F \bullet \theta)$ for each $\theta : G \leftarrow H$,
- (c) $\text{zip}.F.(G \bullet H) = (G \bullet \text{zip}.F.H) \circ (\text{zip}.F.G \bullet H)$ for all G and H ,
- (d) $\text{zip}.F.\text{id} = \text{id} \bullet F$.

□

Definition 2 (Commuting Relators) The half-zip $\text{zip}.F.G$ is said to be a *zip* of (F, G) if there exists a half-zip $\text{zip}.G.F$ such that

$$\text{zip}.F.G = (\text{zip}.G.F)^\cup$$

We say that datatypes F and G *commute* if there exists a zip for (F, G) .

□

Constructing Zips

See Hoogendijk's thesis for how these are calculated:

$$\begin{aligned}
 \text{zip}.K_A.G &= \text{fan}.G \bullet K_A \ , \\
 \text{zip}.+.G &= G.\text{inl} \nabla G.\text{inr} \ , \\
 \text{zip}.\times.G &= (G.\text{outl} \triangle G.\text{outr})^\cup \ , \\
 \text{zip}.T.G &= ([\text{id}_G \otimes ; G.\text{in} \circ (\text{zip}.\otimes.G \bullet \text{Id} \triangle T)]) \ .
 \end{aligned}$$

where T is the tree relator with pattern relator $\otimes$.

$$\begin{aligned}
 \text{fan}.K_A &= \top\top_{A,-} \\
 \text{fan}.+ &= (\text{id} \nabla \text{id})^\cup \\
 \text{fan}.\times &= \text{id} \triangle \text{id} \\
 \text{fan}.T &= ([\text{id} \otimes ; (\text{fan}.\otimes)^\cup])^\cup
 \end{aligned}$$

where T is the tree relator with pattern relator $\otimes$.