Constrained datatypes, invariants and business rules: a relational approach

JNO, DI/UM
jno@di.uminho.pt
PURe Project (POSI/CHS/44304/2002)

PUReCafé - May 20th, 2004

Constraints: business rules, invariants

- Banking operation:

$$debit(amount) \quad : \quad Account \longrightarrow Account$$

  Business rule: *account balance allways greater than negotiated minimum.*

- List operation:

$$merge \quad : \quad A^\star \times A^\star \longrightarrow A^\star$$

  Invariant: *lists must be sorted.*

Business rules = (commercial) invariants

Respect for business rules (invariants)

Standard approach: first you have to *invent*...

```
merge (l,[])                  = l
merge ([],r)                  = r
merge (x:xs,y:ys) | x < y     = x : merge(xs,y:ys)
                  | otherwise = y : merge(x:xs,ys)
```

and then *verify*:

$$sorted(merge(l,r)) \quad \Leftarrow \quad (sorted\ l) \wedge (sorted\ r)$$

(Pointwise proofs, theorem provers, etc)

"Respect by construction"

Alternative approach: given

$$A \xleftarrow{f} A$$

and invariant $Bool \xleftarrow{\phi} A$ to be respected,

> Either you find no way to build f or, if you do, $\phi(f\ a) \Leftarrow (\phi\ a)$ is ensured *by construction*.

(MPC = mathematics of program **construction**)

> Constructive proofs: pointfree calculation in the relational calculus

Relations which ensure properties

For any $B \xleftarrow{R} A$ and property $2 \xleftarrow{\phi} B$, R will ensure ϕ iff

$$bRa \quad \Rightarrow \quad \phi\, b$$

It is *always* possible find some ψ such that ψ-*pre-conditioned* R ensures ϕ:

$$bRa \wedge \psi\, a \quad \Rightarrow \quad (\phi\, b) \tag{1}$$

that is (introduce coreflexives $\Psi = [\![\psi]\!]$, $\Phi = [\![\phi]\!]$):

$$\mathit{rng}\,(R \cdot \Psi) \quad \subseteq \quad \Phi \tag{2}$$

Why is (2) "better" than (1)?

Predicates (invariants, etc) are coreflexives

Strategy: identify every

- **predicate** $A \xrightarrow{\phi} \mathit{bool}$ with binary relation $[\![\phi]\!]$ such that $a[\![\phi]\!]b \equiv a = b \wedge (\phi\ a)$.
- So $[\![\phi]\!]$ is **coreflexive**: $[\![\phi]\!] \subseteq id$, cf.

Predicates	Coreflexives
$\lambda x.\ \mathrm{T}$	id
$\lambda x.\ \mathrm{F}$	$\bot$
$p \vee q$	$[\![p]\!] \cup [\![q]\!]$
$p \wedge q$	$[\![p]\!] \cdot [\![q]\!]$
$\neg p$	$id - [\![p]\!]$

“GaloisCalc”

Since *rng* and $(R\cdot)$ can be found as lower adjoints in

$(f\ X) \subseteq Y \equiv X \subseteq (g\ Y)$			
Description	$f = g^\flat$	$g = f^\sharp$	**Obs.**
Left-division	$(R\cdot)$	$(R\setminus\ )$	read “R under ...”
Range	*rng*	$(\cdot\top)$	$\top = !^\circ \cdot !$ and lower $\subseteq$ is restricted to coreflexives
Definitions	$\begin{array}{rcl} f\ X &=& \bigcap\{Y \mid X \subseteq gY\} \\ g\ Y &=& \bigcup\{X \mid f\ X \subseteq Y\}\end{array}$		

(3)

we get (for free):

$$\mathit{rng}\,(R\cdot X) \subseteq Y \quad \equiv \quad X \subseteq g_R\ Y$$

What is the upper adjoint g_R?

Weakest liberal pre-conditions

$g_R\ X$ is well-known — the largest of all pre-conditions over R which ensure X, written $R \mathbin{\blacktriangleright} X$:

$$R \mathbin{\blacktriangleright} Y \quad = \quad \bigcup\{X \mid \mathit{rng}\,(R\cdot X) \subseteq Y\}$$

Alternatively:

$$R \mathbin{\blacktriangleright} Y \quad = \quad \mathit{dom}\,(Y\cdot R) \cup (id - \mathit{dom}\,R)$$

Pointwise version (back to predicates):

$$(R \mathbin{\blacktriangleright} y)a \quad = \quad \forall b \in B.\ bRa \Rightarrow (y\ b)$$

Universal property

We will **never** use any of these definitions. Instead, we resort to universal property

$$\mathit{rng}\,(R \cdot X) \subseteq Y \quad \equiv \quad X \subseteq R \searrow Y \tag{4}$$

that is

$$X \subseteq R \searrow Y \quad \equiv \quad R \cdot X \subseteq Y \cdot R \cdot X \tag{5}$$

($\subseteq$ restricted to coreflexives) or even

$$X \subseteq R \searrow Y \quad \equiv \quad R \cdot X = Y \cdot R \cdot X$$

since $Y \cdot R \subseteq R$.

Galois properties of $R \searrow \Phi$

Conjunction [a]

$$R \searrow (\Phi \cdot \Psi) \quad = \quad (R \searrow \Phi) \cdot (R \searrow \Psi) \tag{6}$$

Reflexion [b]

$$R \searrow Y = id \text{ wherever } \mathit{rng}\, R \subseteq Y$$

Composition [c]

$$(S \cdot R) \searrow \Phi \quad \equiv \quad R \searrow (S \searrow \Phi) \tag{7}$$

[a]=Upper-adjoint distributivity.
[b]$X = id$ is unit of composition (Galois + monoids).
[c]Galois + monoids.

Galois properties of $R \searrow \Phi$

Cancellation

$$\mathit{rng}\,(R \cdot R \searrow Y) \subseteq Y \tag{8}$$

ie

$$R \cdot (R \searrow Y) \subseteq Y \cdot R \cdot (R \searrow Y) \tag{9}$$

entailing $R \cdot (R \searrow Y) \subseteq Y \cdot R$.

More properties

Functions

$$f \mathbin{\diagdown\!\!\!\bullet} \Phi = (f^{\circ} \cdot \Phi \cdot f) \cap id \tag{10}$$

Particular identities

$$\begin{aligned}
id \mathbin{\diagdown\!\!\!\bullet} \Phi &= \Phi && (11)\\
R \mathbin{\diagdown\!\!\!\bullet} id &= id && (12)\\
\bot \mathbin{\diagdown\!\!\!\bullet} \Phi &= id && (13)\\
(id - \mathsf{dom}\, R) &\subseteq R \mathbin{\diagdown\!\!\!\bullet} \Phi && (14)
\end{aligned}$$

etc.

Constrained datatypes

Prospect of category whose objects are coreflexives Ψ, Φ (constraints) and whose arrows are relations which ensure such constraints, that is, every $\Psi \xleftarrow{R} \Phi$ is — by **construction** — such that

$$\Phi \subseteq R \mathbin{\diagdown\!\!\!\bullet} \Psi \tag{15}$$

Check composition: $\Gamma \xleftarrow{S} \Psi$ and $\Psi \xleftarrow{R} \Phi$ compose relationally, yielding $\Gamma \xleftarrow{S \cdot R} \Phi$

$$\frac{\begin{array}{lcl}\Psi & \subseteq & S \mathbin{\diagdown\!\!\!\bullet} \Gamma\\ \Phi & \subseteq & R \mathbin{\diagdown\!\!\!\bullet} \Psi\end{array}}{\Phi \subseteq (S \cdot R) \mathbin{\diagdown\!\!\!\bullet} \Gamma} \tag{16}$$

Proof of (16)

$$
\begin{array}{ll}
 & \Phi \subseteq R \ltimes \Psi \quad \wedge \quad \Psi \subseteq S \ltimes \Gamma \\
\Rightarrow & \qquad \{ \ R \ltimes \_ \text{ is an upper adjoint, thus monotone } \} \\
 & \Phi \subseteq R \ltimes \Psi \quad \wedge \quad R \ltimes \Psi \subseteq R \ltimes (S \ltimes \Gamma) \\
\Rightarrow & \qquad \{ \ \subseteq\text{-transitivity } \} \\
 & \Phi \subseteq R \ltimes (S \ltimes \Gamma) \\
\equiv & \qquad \{ \ (7) \ \} \\
 & \Phi \subseteq (S \cdot R) \ltimes \Gamma
\end{array}
$$

Constraints as types

- Think of constraints Φ, Ψ as types.
- Type polymorphism: arbitrary R is an inhabitant of type $\Psi \longleftarrow \Phi$ provided (15) holds.
- One always has $id \xleftarrow{R} \Phi$, since $R \ltimes id = id$ and Φ is coreflexive.
- In the "limit", $id \xleftarrow{R} id$ always is a valid type assignment, the most general one (in fact, the "conventional" one).
- Don't we write e. $1 + A \times f$ for $id + id \times f$? Consistent with id (the largest coreflexive of its type) also being the the smalest *equivalence relation* on its carrier type. (Cf. initial algebras).

Basics

Ordering:

$$\frac{\Phi \xleftarrow{R} \Phi \quad , \; S \subseteq R}{\Phi \xleftarrow{S} \Phi}$$

Constraint composition (intersection):

$$\frac{\begin{array}{c}\Phi' \xleftarrow{R} \Phi \\ \Psi' \xleftarrow{R} \Psi\end{array}}{(\Phi' \cdot \Psi') \xleftarrow{R} (\Phi \cdot \Psi)} \tag{17}$$

Basic operators

Union:

$$\frac{\begin{array}{c}\Psi \xleftarrow{R} \Phi \\ \Psi \xleftarrow{S} \Phi\end{array}}{\Psi \xleftarrow{R \cup S} \Phi} \tag{18}$$

Intersection:

$$\frac{\begin{array}{c}\Psi \xleftarrow{R} \Phi \\ \Psi \xleftarrow{S} \Phi\end{array}}{\Psi \xleftarrow{R \cap S} \Phi} \tag{19}$$

Subtypes

Subtype ordering: $\Phi' \subseteq \Phi$ (a complete lattice, $\bot = \emptyset, \top = id$)

- Variance:

$$\frac{\Psi \xleftarrow{R} \Phi \ , \Phi' \subseteq \Phi}{\Psi \xleftarrow{R} \Phi'} \tag{20}$$

- Contravariance:

$$\frac{\Psi' \subseteq \Psi, \ \Psi' \xleftarrow{R} \Phi}{\Psi \xleftarrow{R} \Phi} \tag{21}$$

Invariant preservation

Pointwise:

$$a' R a \wedge \phi\, a \Rightarrow \phi\, a'$$

Pointfree ($\Phi = [\![\phi]\!]$):

$$\mathit{rng}\,(R \cdot \Phi) \subseteq \Phi \quad \equiv \quad \Phi \subseteq R \wr \Phi$$

that is,

$$\Phi \xleftarrow{R} \Phi$$

Moving on towards induction

From

$$\begin{array}{ccc} A & \xrightarrow{f} & \mathsf{F}\,A \\ {\scriptstyle [\![\, g,f \,]\!]}\downarrow & & \downarrow{\scriptstyle \mathsf{F}[\![\, g,f \,]\!]} \\ B & \xleftarrow[g]{} & \mathsf{F}\,B \end{array}$$

to

$$\begin{array}{ccc} \Phi & \xrightarrow{S} & \mathsf{F}\,\Phi \\ {\scriptstyle [\![\, R,S \,]\!]}\downarrow & & \downarrow{\scriptstyle \mathsf{F}[\![\, R,S \,]\!]} \\ \Psi & \xleftarrow[R]{} & \mathsf{F}\,\Psi \end{array}$$

We proceed step by step:

Relators

A **relator** is a functor on relations

$$
\begin{array}{ccc}
A & \cdots\cdots & \mathsf{F}\,A \\
{\scriptstyle R}\downarrow & & \downarrow{\scriptstyle \mathsf{F}\,R} \\
B & \cdots\cdots & \mathsf{F}\,B
\end{array}
$$

which is monotonic and commutes with converse:

$$
\begin{array}{lcl}
R \subseteq S & \Rightarrow & (\mathsf{F}\,R) \subseteq (\mathsf{F}\,S) \\
\mathsf{F}\,(R^\circ) & = & (\mathsf{F}\,R)^\circ
\end{array}
$$

(Recall that F will commute with *composition* and *identity* too.)

Coreflexives are preserved by relators

Easy to check:

$$
\begin{array}{ll}
 & \Phi \subseteq id \\
\Rightarrow & \quad \{ \text{ relator } \} \\
 & \mathsf{F}\,\Phi \subseteq \mathsf{F}\,id \\
\equiv & \quad \{ \text{ relator-id } \} \\
 & \mathsf{F}\,\Phi \subseteq id
\end{array}
$$

Constrained F- "maps"

Easy to infer $\mathsf{F}\,\Psi \xleftarrow{\mathsf{F}\,R} \mathsf{F}\,\Phi$ in

$$
\begin{tikzcd}
\Phi \arrow[r, dotted, no head] \arrow[d, "R"'] & \mathsf{F}\,\Phi \arrow[d, "\mathsf{F}\,R"] \\
\Psi \arrow[r, dotted, no head] & \mathsf{F}\,\Psi
\end{tikzcd}
$$

from $\Psi \xleftarrow{R} \Phi$, as follows:

$$
\begin{array}{ll}
 & \Psi \xleftarrow{R} \Phi \\
\equiv & \qquad \{ \text{ definition } \} \\
 & \Phi \subseteq R \mathbin{\backslash\!\!\bullet} \Psi \\
\equiv & \qquad \{ \text{ universal property (5) } \} \\
 & R \cdot \Phi \subseteq \Psi \cdot R \cdot \Phi \\
\Rightarrow & \qquad \{ \text{ relator } \} \\
 & \mathsf{F}\,R \cdot \mathsf{F}\,\Phi \subseteq \mathsf{F}\,\Psi \cdot \mathsf{F}\,R \cdot \mathsf{F}\,\Phi \\
\equiv & \qquad \{ \text{ universal property (5) } \} \\
 & \mathsf{F}\,\Phi \subseteq \mathsf{F}\,R \mathbin{\backslash\!\!\bullet} \mathsf{F}\,\Psi \\
\equiv & \qquad \{ \text{ definition } \} \\
 & \mathsf{F}\,\Psi \xleftarrow{\mathsf{F}\,R} \mathsf{F}\,\Phi
\end{array}
$$

Constrained F-algebras

Check what it means to write

$$\Phi \xleftarrow{R} \mathsf{F}\,\Phi$$

We get

$$\mathsf{F}\,\Phi \quad \subseteq \quad R \setminus\!\!\bullet\, \Phi$$

that is,

$$R \cdot \mathsf{F}\,\Phi \quad \subseteq \quad \Phi \cdot R \cdot \mathsf{F}\,\Phi$$

Since $\Phi \cdot R \cdot \mathsf{F}\,\Phi \quad \subseteq \quad \Phi \cdot R$ we get, by monotonicity:

$$R \cdot \mathsf{F}\,\Phi \quad \subseteq \quad \Phi \cdot R \tag{22}$$

In other words: Φ is a (coreflexive) F-*congruence* for R.

F-congruences

(See *When is a function a fold or an unfold?* [GHA01])

- **Congruences** are (endo) relations which are preserved along some kind of **algebraic** structure.

  (Term "congruence" is too strong: it might be better to call these *compatible* relations, *cf.* the terminology of [Blo76].)

- Informally, every operation in such a structure applied to congruent arguments should yield congruent results.
- Algebraic structure is captured by the concept of a **relator**.

F-congruence

Given relation $A \xleftarrow{R} \mathsf{F}\,A$ (a so-called F-*algebra*), we say that relation $A \xleftarrow{T} A$ is an F-*congruence for* R iff

$$R \cdot \mathsf{F}\,T \subseteq T \cdot R \qquad \begin{tikzcd} A \arrow[d, "T"'] \arrow[dr, phantom, "\supseteq"] & \mathsf{F}\,A \arrow[l, "R"'] \arrow[d, "\mathsf{F}\,T"] \\ A & \mathsf{F}\,A \arrow[l, "R"] \end{tikzcd} \tag{23}$$

Example

$A = A^\star$, $R = [nil, cons]$ and $T = \leq$ ("prefix"):

$$\begin{tikzcd} nil \arrow[d, mapsto, "\leq"'] & cons(a,l) \arrow[d, mapsto, "\leq"'] \arrow[dr, phantom, "\supseteq"] & (a,l) \arrow[l, mapsto, "cons"'] \arrow[d, mapsto, "id\times\leq"] \\ nil & cons(a,l') & (a,l') \arrow[l, mapsto, "cons"] \end{tikzcd}$$

that is:

$$nil \leq nil$$
$$l' \leq l \Rightarrow cons(a,l') \leq cons(a,l)$$

Constrained catamorphisms

Checking the consistency of diagram

$$\begin{tikzcd} \mu\mathsf{F} \arrow[d, "(|R|)"'] & \mathsf{F}(\mu\mathsf{F}) \arrow[l, "in"'] \arrow[d, "\mathsf{F}(|R|)"] \\ \Phi & \mathsf{F}\,\Phi \arrow[l, "R"] \end{tikzcd}$$

First, we need to know about (two forms of) relational cata-fusion [BdM97]:

$$(|T|) \subseteq S \cdot (|R|) \quad \Leftarrow \quad T \cdot \mathsf{F}\,S \subseteq S \cdot R \tag{24}$$
$$S \cdot (|R|) \subseteq (|T|) \quad \Leftarrow \quad S \cdot R \subseteq T \cdot \mathsf{F}\,S \tag{25}$$

Checking the correctness of the constrained $(\!|R|\!)$

$$\Phi \xleftarrow{(\!|R|\!)} \mu\mathsf{F}$$

$$\equiv \quad \{ \text{ definition } \}$$

$$id \quad \subseteq \quad (\!|R|\!) \mathbin{\bullet\!\!\!\backslash} \Phi$$

$$\equiv \quad \{ \text{ definition } \}$$

$$(\!|R|\!) \quad \subseteq \quad \Phi \cdot (\!|R|\!)$$

$$\Leftarrow \quad \{ \text{ relational cata-fusion } \}$$

$$R \cdot \mathsf{F}\,\Phi \quad \subseteq \quad \Phi \cdot R$$

$$\Leftarrow \quad \{ \text{ (22) above } \}$$

$$\Phi \xleftarrow{R} \mathsf{F}\,\Phi$$

Example — Sorting

Given

$$IsSorted \stackrel{\mathrm{def}}{=} (\!|in \cdot (id + ok)|\!)$$

for

$$ok(a, x) = \forall b \in \textit{elems}\, x.\ a \leq b$$

build

$$\begin{tikzcd}
A^\star \arrow[d, "(\!|R|\!)"'] & 1 + A \times A^\star \arrow[l, "in"'] \arrow[d, "\mathsf{F}\,(\!|R|\!)"] \\
IsSorted & 1 + A \times IsSorted \arrow[l, "R"]
\end{tikzcd}$$

Sorting continued

R must be such that

$$\begin{tikzcd} A^\star \arrow[d, "IsSorted"'] & & 1 + A \times A^\star \arrow[ll, "R"'] \arrow[d, "1+A\times IsSorted"] \\ A^\star & \supseteq & 1 + A \times A^\star \arrow[ll, "R"] \end{tikzcd}$$

$$R \cdot (1 + A \times IsSorted) \quad \subseteq \quad IsSorted \cdot R$$

Sorting continued

$$\begin{array}{ll}
& R \cdot (1 + A \times IsSorted) \quad \subseteq \quad IsSorted \cdot R \\
\equiv & \quad \{ \text{ expanding and restricting to functions } \} \\
& [r_1, r_2] \cdot (1 + A \times IsSorted) \quad \subseteq \quad IsSorted \cdot [r_1, r_2] \\
\Leftarrow & \quad \{ \text{ expanding } \} \\
& \begin{array}{rcl} r_1 & \subseteq & IsSorted \cdot r_1 \\ r_2 \cdot (id \times IsSorted) & \subseteq & IsSorted \cdot r_2 \end{array} \\
\Leftarrow & \quad \{ \text{ shunting (Galois connections over functions) } \} \\
& \begin{array}{rcl} \mathit{img}\, r_1 & \subseteq & IsSorted \\ (id \times IsSorted) & \subseteq & r_2^\circ \cdot IsSorted \cdot r_2 \end{array} \\
\equiv & \quad \{ \text{ choose } r_1 = nil = \underline{[\,]} \} \\
& (id \times IsSorted) \quad \subseteq \quad r_2^\circ \cdot IsSorted \cdot r_2 \\
\equiv & \quad \{ \text{ .... } \} \\
& \ldots\ldots
\end{array}$$

etc. “à la” Bird-Moor [BdM97].

Insertion sort

I haven't checked, but we should be able to find solution $r_2 = insert$, where

```
insert(x,[])                = [x]
insert(x,a:l) | x < a       = [x,a]++l
              | otherwise   = a:(insert(x,l))
```

Comments:

- Still a lot to be done
- Constrained hylos, constrained F-coalgebras, etc

Constrained F-coalgebras

Check what it means to write

$$\Phi \xrightarrow{R} \mathsf{F}\,\Phi$$

We get

$$\Phi \quad \subseteq \quad R \mathbin{\backslash\!\!\!\bullet} (\mathsf{F}\,\Phi)$$

that is,

$$R \cdot \Phi \quad \subseteq \quad (\mathsf{F}\,\Phi) \cdot R \cdot \Phi$$

Since $(\mathsf{F}\,\Phi) \cdot R \cdot \Phi \quad \subseteq \quad (\mathsf{F}\,\Phi) \cdot R$ we get, by monotonicity:

$$R \cdot \Phi \quad \subseteq \quad \mathsf{F}\,\Phi \cdot R \tag{26}$$

In other words: Φ is a (coreflexive) F-*invariant* for R.

F-invariants

(See *When is a function a fold or an unfold?* [GHA01])

Duality: given relation $\mathsf{F}\,A \xleftarrow{S} A$ (a so-called F-*coalgebra*), we say that relation $A \xleftarrow{R} A$ is an F-*invariant for S* iff

$$S \cdot R \;\subseteq\; \mathsf{F}\,R \cdot S \qquad \begin{tikzcd} A \arrow[r, "S"] & \mathsf{F}\,A \\ A \arrow[u, "R"] \arrow[r, "S"'] \arrow[ur, phantom, "\supseteq"] & \mathsf{F}\,A \arrow[u, "\mathsf{F}\,R"'] \end{tikzcd} \tag{27}$$

Further work — invariant refinement

Express the "SETS" laws [Oli92] as invariant-refinements rather than datatype refinements:

- Say that $\Psi <_R \Phi$ wherever ...... etc
  Example: $IsSorted <_{(ff2seq^\circ)} Monotone$
- Say that $\Psi \cong_R \Phi$ wherever .... etc
  Example: $id \cong_{(join^\circ)} Eqdom$ will replace

$$A \rightharpoonup (B \times C) \leq (A \rightharpoonup B) \times (A \rightharpoonup C) \tag{28}$$

Cf. PhD work by C. Rodrigues.

References

[BdM97] R. Bird and O. de Moor. *Algebra of Programming*. Series in Computer Science. Prentice-Hall International, 1997. C. A. R. Hoare, series editor.

[Blo76] S.L. Bloom. Varieties of ordered algebras. *JCSS*, 13:200–212, 1976.

[GHA01] Jeremy Gibbons, Graham Hutton, and Thorsten Altenkirch. When is a function a fold or an unfold? *Electronic Notes in Theoretical Computer Science*, 44(1), 2001.

[Oli92] J. N. Oliveira. *Software Reification using the SETS Calculus* . In Tim Denvir, Cliff B. Jones, and Roger C. Shaw, editors, *Proc. of the BCS FACS 5th Refinement Workshop, Theory and Practice of Formal Software Development, London, UK*, pages 140–171. ISBN 0387197524, Springer-Verlag, 8–10 January 1992. (Invited paper).