InvarGenT: Implementation

BY ŁUKASZ STAFINIAK

Institute of Computer Science
University of Wrocław

Abstract

InvarGenT is a proof-of-concept system for invariant generation by full type inference with Guarded Algebraic Data Types and existential types encoded as automatically generated GADTs. This implementation documentation focuses on source code, refers to separate technical reports on theory and algorithms.

1 Data Structures and Concrete Syntax

Following [1], we have the following nodes in the abstract syntax of patterns and expressions:

`p-Empty.` 0: Pattern that never matches. Concrete syntax: `!`. Constructor: `Zero`.

`p-Wild.` 1: Pattern that always matches. Concrete syntax: `_`. Constructor: `One`.

`p-Var.` x: Pattern variable. Concrete syntax: lower-case identifier e.g. `x`. Constructor: `PVar`.

`p-Cstr.` $Kp_1 ... p_n$: Constructor pattern. Concrete syntax: e.g. `K (p1, p2)`. Constructor: `PCons`.

`Var.` x: Variable. Concrete syntax: lower-case identifier e.g. `x`. Constructor: `Var`. External functions are represented as variables in global environment.

`Cstr.` $Ke_1 ... e_n$: Constructor expression. Concrete syntax: e.g. `K (e1, e2)`. Constructor: `Cons`.

`App.` $e_1\, e_2$: Application. Concrete syntax: e.g. `x y`. Constructor: `App`.

`LetRec.` $\mathbf{letrec}\, x = e_1\, \mathbf{in}\, e_2$: Recursive definition. Concrete syntax: e.g. `let rec f = function ... in ...` Constructor: `Letrec`.

`Abs.` $\lambda(c_1 ... c_n)$: Function defined by cases. In concrete syntax only implemented for single branching via `fun` keyword, e.g. `fun x y -> f x y` translates as $\lambda(x.\lambda(y.(fx)\ y))$. Constructor: `Lam` (allows multiple branches).

`Clause.` $p.e$: Branch of pattern matching. Concrete syntax: e.g. `p -> e`.

`CstrIntro.` Does not figure in neither concrete nor abstract syntax. Scope of existential types is thought to retroactively cover the whole program.

`ExCases.` $\lambda[K](p_1.e_1 ... p_n.e_n)$: Function defined by cases and abstracting over the type of result. Concrete syntax: e.g. `function Nil -> ... | Cons (x,xs) -> ...` Parsing introduces a fresh identifier for K. Constructor: `ExLam`.

`ExLetIn.` $\mathbf{let}\ p = e_1\ \mathbf{in}\ e_2$: Elimination of existentially quantified type. Concrete syntax: e.g. `let v = f e ... in ...` Constructor: `Letin`.

For types and formulas, we have ASCII and unicode syntactic variants (the difference is only in lexer). Quantified variables can be space or comma separated. The table below is analogous to information for expressions above. Existential type construct introduces a fresh identifier for K.

type variable	x	`x`	`x`	`TVar`
type constructor	List	`List`	`List`	`TCons`
number	7	`7`	`7`	`NCst`
existential type	$\exists\alpha\beta[a \leqslant \beta].\tau$	`ex a b [a<=b].t`	`∃a,b[a≤b].t`	`TExCons`
type sort	s_{ty}	`type`		
number sort	s_R	`num`		
function type	$\tau_1 \to \tau_2$	`t1 -> t2`	`t1 → t2`	`Fun`
equation	$a \doteq b$	`a = b`	`a = b`	`Eqty`
inequation	$a \leqslant b$	`a <= b`	`a ≤ b`	`Leq`
conjunction	$\varphi_1 \wedge \varphi_2$	`a=b && b=a`	`a=b ∧ b=a`	built-in lists

Toplevel expressions (corresponding to structure items in OCaml) introduce types, type and value constructors, global variables with given type (external names) or inferred type (definitions).

type constructor	`newtype List : type * num`	`TypConstr`
value constructor	`newcons Cons : all n a. a * List(a,n) --> List(a,n+1)`	`ValConstr`
	`newcons Cons : ∀n,a. a * List(a,n) ⟶ List(a,n+1)`	
declaration	`external filter : ∀n,a. List(a,n)→∃k[k<=n].List(a,k)`	`PrimVal`
rec. definition	`let rec f =...`	`LetRecVal`
non-rec. definition	`let v =...`	`LetVal`

Mutual non-nested recursion and or-patterns are preserved by parsing but will be eliminated by source-code transformation afterwards.

Bibliography

[1] Łukasz Stafiniak. A gadt system for invariant inference. Manuscript, 2012. Available at: http://www.ii.uni.wroc.pl/~lukstafi/pubs/EGADTs.pdf