Logic programming and types: An introduction to Mercury

Sergei Winitzki

SF Types, Theorems, and Programming Languages meetup

April 18, 2016

Logic programming as a DSL for logic puzzles (Prolog)

All jumping creatures are green. All small jumping creatures are martians.
All green martians are intelligent.
Ngtrks is small and green. Pgvdrk is a jumping martian.
Who is intelligent? (inpired by S. Lem, *Invasion from Aldebaran*)

```
small(ngtrks). green(ngtrks).
martian(pgvdrk). jumping(pgvdrk).
green(X) :- jumping(X).
martian(X) :- small(X), jumping(X).
intelligent(X) :- green(X), martian(X).
?- intelligent(X).
X = pgvdrk
```

- Classical logic limited to Horn clauses = executable program
- The legacy of Prolog: SQL, Datalog, ... [_ | X | Xs]

Logic programming refresher

- All capitalized symbols are **quantified logical variables** (LVs)
- LVs on the left imply $\forall$; free LVs on the right imply $\exists$
- The symbol `:-` means implication $\leftarrow$, the comma means "*and*"
- Rules can be given in any order; no declarations
- Examples:

  `green(X) :- jumping(X)` means $\forall X : \text{jumping}(X) \rightarrow \text{green}(X)$

  `path(A,B) :- path(A,C), path(C,B)` means

  $\forall A, B : [\exists C : \mathit{path}(A, C) \wedge \mathit{path}(C, B) \rightarrow \mathit{path}(A, B)]$
- `?- intelligent(X)` means: prove that $\exists$`X:intelligent(X)`
- The runtime will perform a backtracking search for a proof
- LVs will be assigned/unassigned as needed during search

Mercury in a nutshell: a Prolog dialect with static types

- Logical variables, predicates, rules in Prolog-style syntax and semantics
- Functions as a primitive, on par with predicates
- Higher-order functions and higher-order predicates
- Data structures: list, array, hash map; some support for *mutable* data
- Flexible mode system, type checking, mode and type inference
- Polymorphic types and type classes
- Universal and existential type quantifiers
- Termination and determinism analysis for predicates
- Multithreading; parallel goals; exceptions
- Compilation to C, Java, Erlang, .NET
  - Interop and type-level compatibility with these environments

Modes, determinism, types, I/O

- Predicate arguments have **types**, **modes**, and **determinism**
- "Martians" requires a declaration of modes & determinism for `main`:

```
pred main(io::di, io::uo) is cc_multi.
```

- Input/output: Special syntax `!IO` means a pair of `io` ("world") values
- Proof search may backgrack - the goal `intelligent(X)` may fail, so:

```
main(!IO) :- ( if intelligent(X) then write(X, !IO)
  else write_string("No solution", !IO) ).
```

- Determinism: `det`, `semidet`, `multi`, `nondet`, `failure`, `cc_multi`, `cc_nondet`
- Predefined modes: `in`, `out`, `di`, `uo`
  - A **mode** describes what happens to an LV during proof search:

    ```
    :- mode out == (free >> ground).
    :- mode uo == (free >> unique).
    :- mode di == (unique >> dead).
    ```

  - User-defined modes are possible

Example: integer factorial

- Implementation as a predicate:

```
:- pred fact(int::in, int::out) is det.
fact(N,F) :- ( if N =< 1 then F = 1 else fact(N-1, A), F = A*N ).
```

- Implementation as a function:

```
:- func fct(int) = int.
fct(N) = ( if N =< 1 then 1 else N*fct(N-1) ).
```

- The generated code is *identical* in both cases!
  - However, predicates are **not** “functions returning Boolean”
- Type/mode/deteterminism are inferred and statically checked
- *Order of goals* is inferred from modes rather than from code!

```
fact1(N,F) :- ( if N =< 1 then F = 1 else fact(N-1, A), F = A*N ).
fact2(N,F) :- ( if N =< 1 then F = 1 else F = A*N, fact(N-1, A) ).
```

Static determinism analysis

- Compilation fails if we define the predicate by writing two clauses:

```
:- pred fact(int::in, int::out) is det.
fact(1,1).
fact(N,F) :- fact(N-1, A), F = A*N.
```

- Error message:

```
In ‘fact’(in, out): error: determinism declaration not satisfied.
Declared ‘det’, inferred ‘multi’.
Disjunction has multiple clauses with solutions.
```

- Similar error when defining a *function* by clauses:

```
Error: invalid determinism for ‘fct’(in) = out:
the primary mode of a function cannot be ‘multi’.
```

- This is why we need to use a less readable `if-then-else` construct

Functions and expressions

- Functions are deterministic predicates with "in-out" mode
- Expressions are syntactic sugar for goals involving functions
- Hence, the general syntax for functions:
  `fname(Arg1, Arg2, ...)  = Result :- goal, goal, ..., Result = ...`
- "Lambda" predicate terms and function terms:

```
F1 = ( pred(N::in, X::out) is det :- fact(N, X) ).
F2 = fact . % same value as F1 by η-equivalence
F3 = ( func(N) = X :-  X = fct(N) ).
F4 = fct.   % same value as F3 by η-equivalence
```

- Higher-order predicates:
  `map(fact, [1,3,5], Res)` will unify `Res` with `[1,6,120]`
  - Declaring higher-order predicate modes is verbose and complicated

Algebraic types and polymorphism

- Primitive types: `bool`, `int`, `float`, `string`
  - No "symbol literals" as in Prolog
  - Use union type instead, to make "martians" type-safe! (demo)
- Tuples, unions

```
:- type mytype3 ---> point({int,int},string); ok(string); failed.
:- type list(T) ---> []; [ T | list(T) ]. % special syntax
:- type tree(T) ---> leaf; branch(tree(T), T, tree(T)).
```

- Records with accessors

```
:- type employee ---> employee(name :: string, id :: int).
X = employee(...), if X ^ name = "myself" then ...
```

- Predicate types, function types
  - Type syntax: `pred(int::in, int::out)` and `func(int) = int`

Example: "reversible computation"

- Convert integers between unary encoding and native representation

```
:- type unary ---> z; s(unary). %   z;  s(z);  s(s(z));  etc.
:- pred unary_int(unary, int).
:- mode unary_int(in, out). % is det.
:- mode unary_int(out, in). % is multi.
unary_int(C, N) :- (
  C = z, N = 0 ;
  C = s(B), N = M+1, unary_int(B, M)
% goals will be reordered here depending on mode!
).
```

- The same code will unify `unary_int(s(s(z)), N)` or `unary_int(C, 3)`
  - Determinism and mode of predicates is inferred from use!
  - *Different code* is compiled for each declared mode of `unary_int(C, N)`
- All predicates are functions in disguise

Type classes à la Haskell

- A type class defines a set of predicates and functions

```
:- typeclass showable(T) where [
  pred show(T::in, io::di, io::uo) is det
].
:- instance showable(int) where [
  pred(show/3) is io.write_int
].
:- instance showable(list(T)) <= showable(T) where [
  pred(show/3) is show_list
].
:- pred show_list(list(T), io, io) <= showable(T).
show_list(L, !IO) :- ... % define it now
```

- The methods of a type class are resolved at compile time
  - Only one instance declaration per type (as in Haskell)

No higher-kinded types

- Can we define a "functor" or "monad" typeclass? (No!)
  - Type classes must be parameterized by simple types
  - Instances are for concrete types or concrete type constructors
- Cannot parameterize a type class by an unknown type constructor

```
:- typeclass functor(A,B, FA,FB) where [
  func fmap(func(A)=B) = (func(FA)=FB)
].
:- instance functor(A, B, list(A), list(B)) % fails to compile!
where [
   fmap(F::(func(A)=B)) = (func(L) = FL :- FL = map(F, L))
].
```

Limitations of Mercury

- No higher-kind type constructors
  - (e.g. cannot define "functor" or "monad" type classes)
- No row polymorphism for records
  - (ad-hoc record accessors are provided)
- No partially instantiated values (e.g. no difference lists)
  - Queues and other data types are provided by standard library
- No "cut" operator; determinism is controlled more explicitly
- No REPL or interactive queries

Summary

- How to unite "logic programming" and "functional programming"
- Mercury = Prolog + modes + types
- Functions are predicates, and predicates are functions!
- For more info on Mercury:
  - Features of the Mercury programming language (online link)