An Introduction to the **Flix** Programming Language

Dahl-Nygaard Prize Talk

MAGNUS MADSEN

Thanks to our sponsors:











Magnus



Jonathan



Ondřej



Matt



Jaco



Jakob



Nina

Outline

An introduction to the Flix programming language:

- 1 First-class Datalog
- 2 Effect system
- 3 Local mutation
- 4 Purity reflection
- 5 Design principles
- 6 Ecosystem and tooling

The Flix Principle

In Flix the primary building block is a **function**. A function maps an input to an output.

Flix allows functions to be written in the most natural and/or efficient style ...

- Functionally (i.e. with immutable data structures)
- Imperatively (i.e. with mutable data structures)
- Declaratively (i.e. as a collection of logic constraints)

... without revealing these implementation to the clients.



First-class Datalog

Introduction to Datalog

Datalog is a simple, yet surprisingly powerful declarative logic programming language.

A bit like SQL but with recursion.

Datalog has several important properties:

- 1. Every Datalog program eventually terminates.
- 2. Every Datalog program has a unique solution.
- 3. Datalog has efficient and parallel evaluation strategies.

Datalog has successfully been used in a wide-variety of applications:

- e.g., program analysis, security analysis, data analytics, bio-informatics.
- Datalog has often been used for very specialized applications that have high complexity and/or performance requirements.

Introduction to Datalog (cont'd)

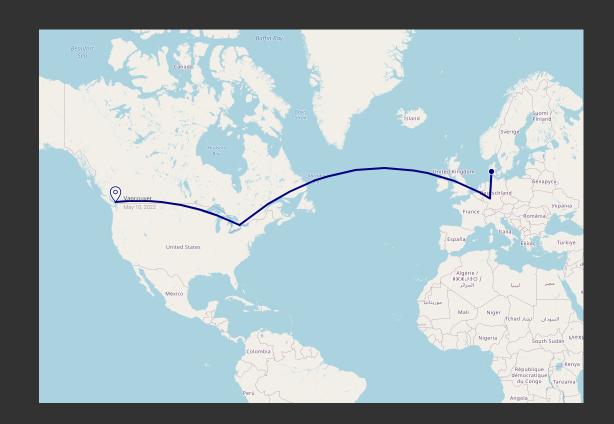
A simple query over two relations:

```
Bird("Eagle").
Bird("Gentoo").
Flying("Eagle").
Penguin(x) :- Bird(x), not Flying(x).
```

A recursive query over two relations:

```
Direct("BLL", "FRA").
Direct("FRA", "YYZ").
Direct("YYZ", "YVR").

Connected(src, dst) :- Direct(src, dst).
Connected(src, dst) :-
   Connected(src, hop), Direct(hop, dst).
```



Motivation

The Problem: Using Datalog is cumbersome!

- I don't want to write my entire program in Datalog.
- I don't want to generate a text file, pass it to a Datalog solver, and parse the result.
- I don't want to use FFI for a 10-line program.
- I want all the conveniences of a real programming language (a type system, an IDE, tools, etc.)
- I want to write *modular* and *reusable* programs.

Key Idea:

Datalog programs as *first-class* values in a functional programming language.

Example: Datalog in Flix

```
def connected(routes: List[(String, String)]): List[(String, String)] =
   let db = project routes into Direct;
   let pr = #{
        Connected(src, dst) :- Direct(src, dst).
        Connected(src, dst) :- Connected(src, hop), Direct(hop, dst).
   };
   query db, pr select (x, y) from Connected(x, y) |> Array.toList
```

Example: Polymorphic Datalog

```
def reachable(edges: List[(t, t)], src: t, dst: t): Bool =
  let db = project edges into Edge;
  let pr = #{
        Path(x, y) :- Edge(x, y).
        Path(x, z) :- Path(x, y), Edge(y, z).
        Reachable() :- Path(src, dst).
  };
  let result = query db, pr select () from Reachable();
  not Array.isEmpty(result)
```

Example: Lattice Semantics

```
def shortestDist(origin: t, edges: List[(t, Int32, t)]): Map[t, N] =
    let db = project edges into Edge;
    let pr = #{
        Dist(origin; N(0)).
        Dist(y; add(d1, d2)) := Dist(x; d1), Edge(x, d2, y).
    };
    query db, pr select (x, d) from Dist(x; d) > Array.toMap
                                                                               0 T
enum N { case N(Int32) }
instance PartialOrder[N] {
    def lessEqual(x: N, y: N): Bool = ...
                                                                               \infty
// LowerBound, JoinLattice, MeetLattice, ...
                                                          Here N = (\infty, 0, \ge, \min, \max)
```

1 Summary: First-Class Datalog

<u>First-class Datalog program values</u> enable us to use Datalog where it shines: to declaratively express and solve fixed-point constraints.

- We can write functions that internally use Datalog program values.
- We can write modular and reusable Datalog programs using polymorphism.
- The type system ensures that our Datalog programs are consistent and stratified.

Flix supports constraints on relations, but also constraints on lattices.

We can solve even more fixed-point problems, including program analyses and two-player games.



Effect System

Type and Effect System

Flix has a type and effect system based on Hindley-Milner.

• The system supports type classes, higher-kinded types, and complete type inference.

The effect system separates pure, impure, and effect polymorphic expressions.

The effect system is the basis for purity reflection.

Tracking purity has several benefits:

- It enables programmers to know when equational reasoning holds.
- It enables the Flix inliner to make more aggressive, but sound, choices.
- It enables the Flix standard library to know when it is safe to parallelize code (more on this later).

Purity

We can express that a function is *pure*:

Here the implementation of add cannot have any side-effects.

We can also express that a higher-order function requires a *pure* function argument:

Here neither f nor count can have any side-effects.

Impurity

We can also express that a function is impure:

It is a type error to annotate an impure function as pure:

```
def illegal() : Unit \ { } =
    println("I am impure!")
```

Effect Polymorphism

We can express that the effect of a higher-order function depends on its argument:

```
def map(f: a \rightarrow b \setminus ef, l: List[a]): List[b] \ ef = ...

^^ effect variable
```

The effect of map is the same as the effect of f.

```
// Pure use of map
List.map(x -> x * x + 42, 1)
// Impure use of map
List.map(x -> println(x), 1)
```

First-class Functions

We support first-class functions. We can express function composition as:

```
\frac{\mathsf{def} >> (\mathsf{f}: a \rightarrow b \setminus ef1, g: b \rightarrow c \setminus ef2): a \rightarrow c \setminus \{ef1, ef2\}}{} = \dots
```

The effect of the returned function is pure if f and g are pure.

(2) Summary: Effect System

The type and effect system enables us to write pure, impure, and effect polymorphic functions.

We can use the type and effect system to track and enforce purity:

- The Flix standard library enforces that the eq, hash, compare, and toString functions are pure.
- The Flix compiler uses purity information during variable and function inlining.

Related Work:

- Polymorphic effect systems [Lucassen et al., '88]
- Region-based memory management [Tofte et al., '97]
- Programming with algebraic effects and handlers [Plotkin et al., '15]



Local Mutation

Local Mutation

We have seen that Flix tracks purity.

- As soon as a function touches mutable memory it gets tainted with impurity.
- This is cumbersome because impurity then proliferates through the program.
- But what if the use of mutation is in some sense "local".

Can we do better?

- We associate all mutable data with a **region**:
- Reads and writes to data in a region are precisely tracked by the effect system.
- All effects related to a region vanish when the region goes out-of-scope.

Example: Sorting

```
///
/// Sort the given list `l` so that elements are ordered from low to
/// high according to their `Order` instance.

///

def sort(l: List[a]): List[a] with Order[a] =
    region r {
        let arr = List.toArray(l, r);
        Array.sort!(arr);
        Array.toList(arr)
}
1. Introduce a new region.
2. Allocate (mutable) data in the region.
3. Do imperative programming.
4. Return immutable data.
```

We Using an array-based (in place) sort is much faster than any list-based sort.

Example: Adding Two Numbers

```
///
/// Returns the sum of `x` and `y`.
///
def sum(x: Int32, y: Int32): Int32 \ { } =
    region r {
       let a = [x, y] @ r;
       Array.swap!(0, 1, a);
       a[1] + a[0]
    }
```

Example: Swapping Elements

The region is part of the array type.

 \bigcirc The type and effect system tracks reads and writes to regions.

Example: ToString

```
///
/// Returns a String representation of the given list `l`.
///
/// The returned String is of the form x1 :: x2 :: .. :: Nil.
///
def toString(l: List[a]): String with ToString[a] =
    region r {
        let sb = new StringBuilder(r);
        List.foreach(x -> StringBuilder.appendString!("${x} :: ", sb), l);
        StringBuilder.appendString!("Nil", sb);
        StringBuilder.toString(sb)
    }
```

☐ Using StringBuilders in toString functions is intuitive and efficient.

(3) Summary: Local Mutation

Local mutation enables us to write **pure** functions that use **local mutable state**.

- We can implement functions in imperative style.
- We can use imperative style when it is more natural and/or more efficient.
- The type and effect system continues to ensure separation of pure and impure code!
- We can "pretend" to be **functional** programmers but use **imperative** style when we want!

We get the best of both worlds!

Related Work:

- Linear types can change the world! [Wadler et al., '90]
- Imperative functional programming [Jones et al., '93]
- Lazy functional state threads [Jones et al., '94]



Purity Reflection

Purity Reflection

Program evaluation in Flix (and most programming languages) is eager and sequential.

- But it would be useful if library authors could take advantage of lazy and/or parallel evaluation.
- But when is it safe to evaluate a function lazily or in parallel?

Many programming languages have **streams**:

- But the combination of streams and side-effects is a dangerous cocktail.
- We risk race conditions, deadlocks, lost and/or re-ordered side effects!

Can we do better?

 \P What if we allow data structure operations (map, filter, etc.) to vary their behavior depending on the purity of the their function arguments?

Example: Selective Laziness

We can write a map function that uses selective laziness:

```
def map(f: a -> b & ef, l: DelayList[a]): DelayList[b] & ef =
    reifyEff(f) {
        case Pure(g) => mapL(g, l)
        case _ => mapE(f, l)
    }
```

If f is pure then we use mapL to apply it lazily over the list.

If f is impure then we use mapE to apply it eager over the list (materializing all effects).

Example I

The Flix program fragment:

```
DelayList.range(1, 1_000_000_000) |>
DelayList.map(x -> {println("a"); x + 1}) |>
DelayList.map(x -> {println("b"); x * 2})
```

Prints one billion a's followed by one billion b's.

This takes a while, but ultimately the program terminates.

 \bigcirc The a's are printed before the b's preserving the order of effects.

Example II

The Flix program fragment:

```
DelayList.range(1, 1_000_000_000) |>
DelayList.map(x -> x + 1) |>
DelayList.map(x -> x * 2) |>
DelayList.head |> println
```

Prints Some (4) and terminates immediately.

The two map operations are pure, consequently they are applied lazily.

Example III

The Flix program fragment:

```
let count = ref 0;
List.range(1, 1_000_000_000) |>
List.map(x -> x + 1) |>
List.take(1_000) |>
List.map(x -> { count := deref count + 1; x * 2});
println(deref count)
```

Prints 1000 and terminates very quickly.

The first map operation is applied lazily and the subsequent take operation is applied lazily.

The final map operation is applied eagerly, but only to the first 1000 elements.

Example: Selective Parallelism

We can also write a Map function that uses selective parallelism:

```
def mapWithKey(f: (k, v1) -> v2 & ef, t: RBTree[k, v1]): ... =
    reifyEff(f) {
        case Pure(g) => parMapWithKey(g, t)
        case _ => seqMapWithKey(f, t)
}
```

We use this function in the implementation of the **Set** and **Map** data structures.

A Fresh Take on Data Transformations

Principle: Data structure operations (such as map, filter, ...)

- Use <u>lazy</u> and/or <u>parallel</u> <u>evaluation</u> when given <u>pure</u> function arguments.
- Use <u>eager sequential</u> evaluation when given <u>impure</u> function arguments.

This ensures that side-effects are not lost and that the order of side-effects is preserved.

(4) Summary: Purity Reflection

<u>Purity reflection</u> enables higher-order functions to *inspect the purity* of their function argument(s) and to *vary their behavior* based on this information.

We can use this information to implement new and novel data structures:

- DelayList a list that is maximally lazy except when given impure functions.
- DelayMap a map that is lazy in its values and uses parallel evaluation for bulk operations.

Related Work:

First-class effect reflection for effect-guided programming [Long et al., '16]



Design Principles

The Flix Design Principles

A set of **design choices** or **design principles** collected over time.

Based on:

- Discussions on GitHub.
- Discussions on other programming language forums.
- Perceived mistakes of other programming languages.
- Feedback from users.

Listed on the Flix website for reference and to keep us honest.

We have now collected more than 59 such principles.

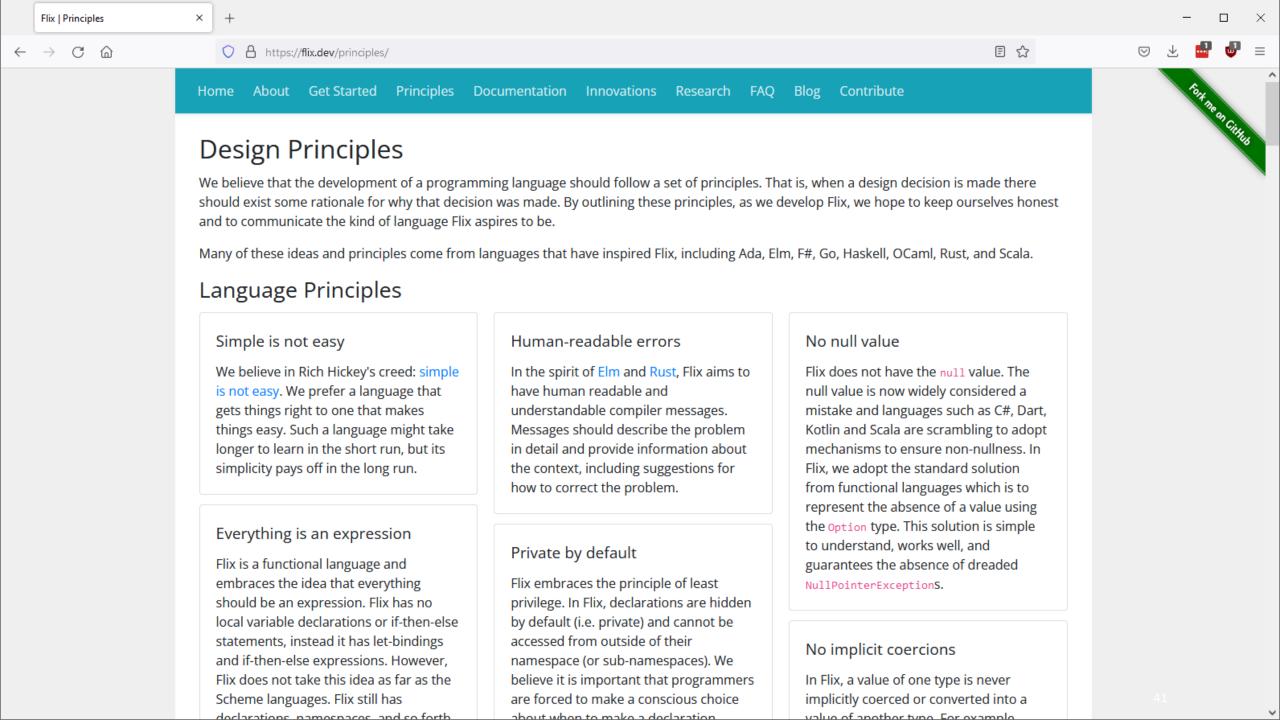
Principle: No warnings, only errors.

Principle: No unused declarations.

Principle: Declaration monotonicity.

```
namespace A {
    def inc(x: Int32): Int32 = x + 1
}
namespace B {
    // ...
}
```

```
def main(): Int32 =
    use A._;
    use B._;
    inc(1)
```





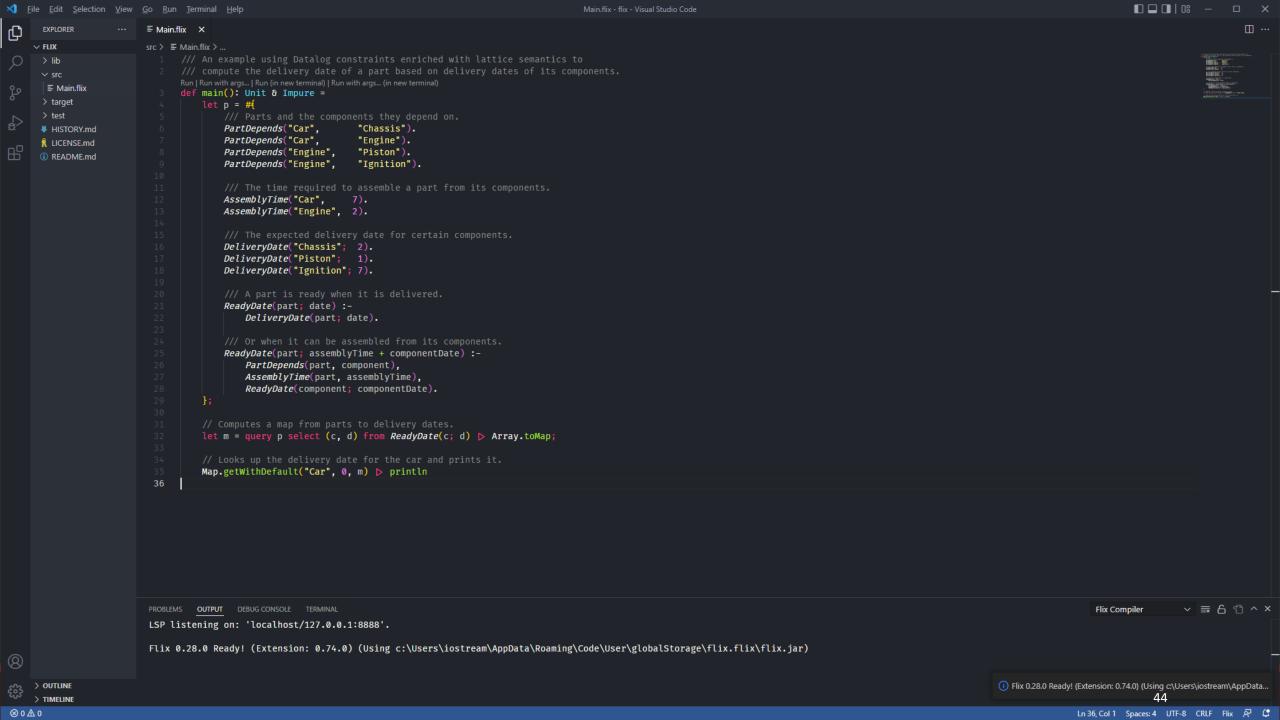
Ecosystem & Tooling

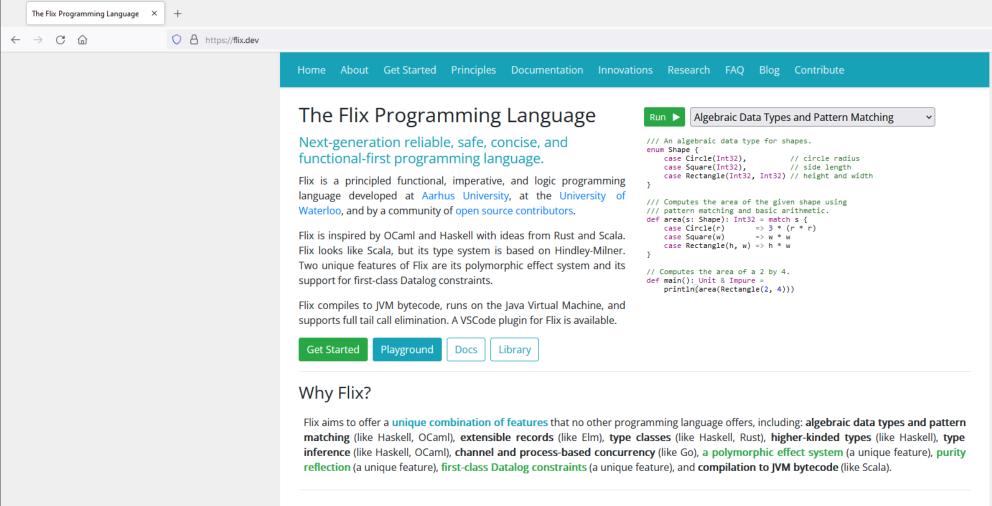
Visual Studio Code Support

We support most Visual Studio Code features, including:

- √ syntax highlighting
- √ inline diagnostics
- √ auto-complete
- √ type and effect hover
- **√** find references
- √ jump to definition
- ✓ rename

- √ code hints
- √ code lenses (e.g. "click to run")
- √ document symbols
- √ workspace symbols
- √ highlight related symbols
- √ incremental compilation





Algebraic Data Types and Pattern Matching

Algebraic data types and pattern matching are the bread-andbutter of functional programming and are supported by Flix with minimal fuss.

```
enum Shape {
    case Circle(Int32),
    case Square(Int32),
    case Rectangle(Int32, Int32)
}

def area(s: Shape): Int32 = match s {
    case Circle(r) => 3 * (r * r)
    case Square(w) => w * w
    case Rectangle(h, w) => h * w
}
```

```
def origin(): (Int32, Int32) = (0, 0)

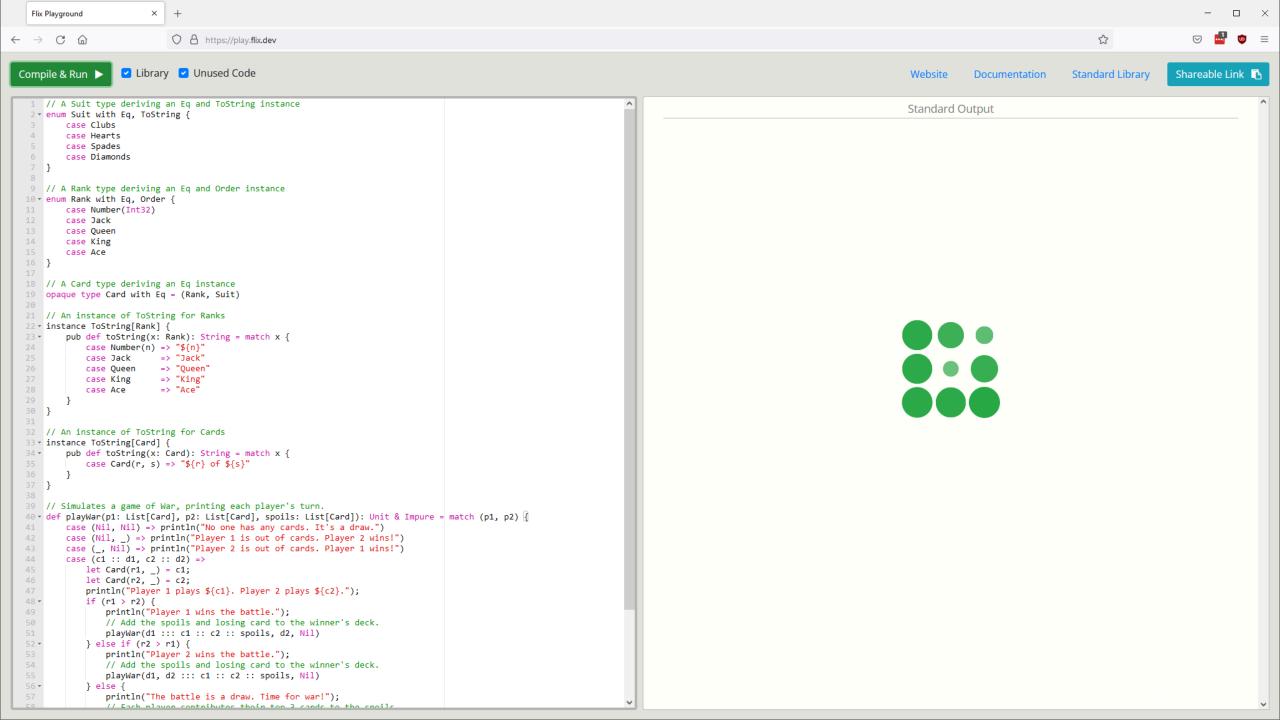
def oneByOne(): {w :: Int32, h :: Int32} = {w = 1, h = 1}

def twoByFour(): {w :: Int32, h :: Int32} = {w = 2, h = 4}
```

Tuples and Records

Flix has built-in support for tuples and records.

☆



Introduction Functions Data Types Lists Records References Arrays Namespaces Concurrency Effects Fixpoints Interoperability Build & Packages Tips & Tricks

Fixpoints

A unique feature of Flix is its built-in support for fixpoint computations on *constraint on relations* and *constraint on lattices*.

We assume that the reader is already familiar with Datalog and focus on the Flix specific features.

Using Flix to Solve Constraints on Relations

We can use Flix to solve a fixpoint computation inside a function.

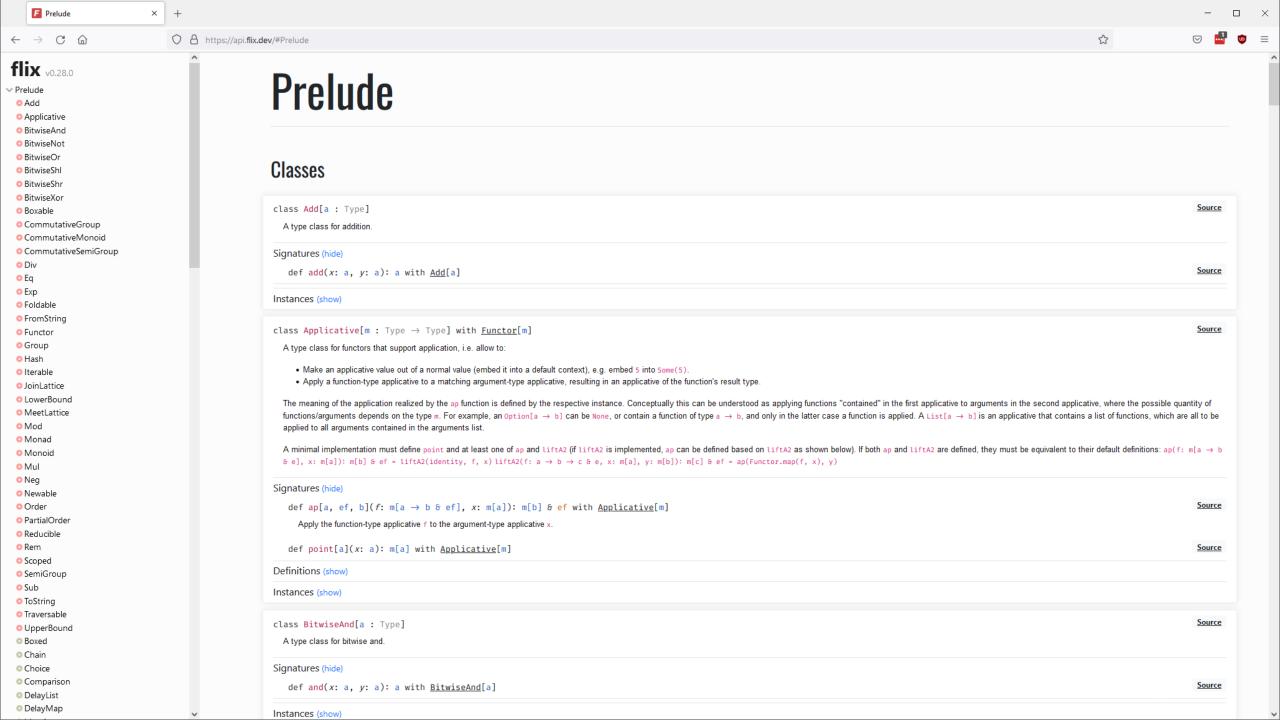
For example, given a set of edges s, a src node, and dst node, compute if there is a path from src to dst. We can elegantly solve this problem as follows:

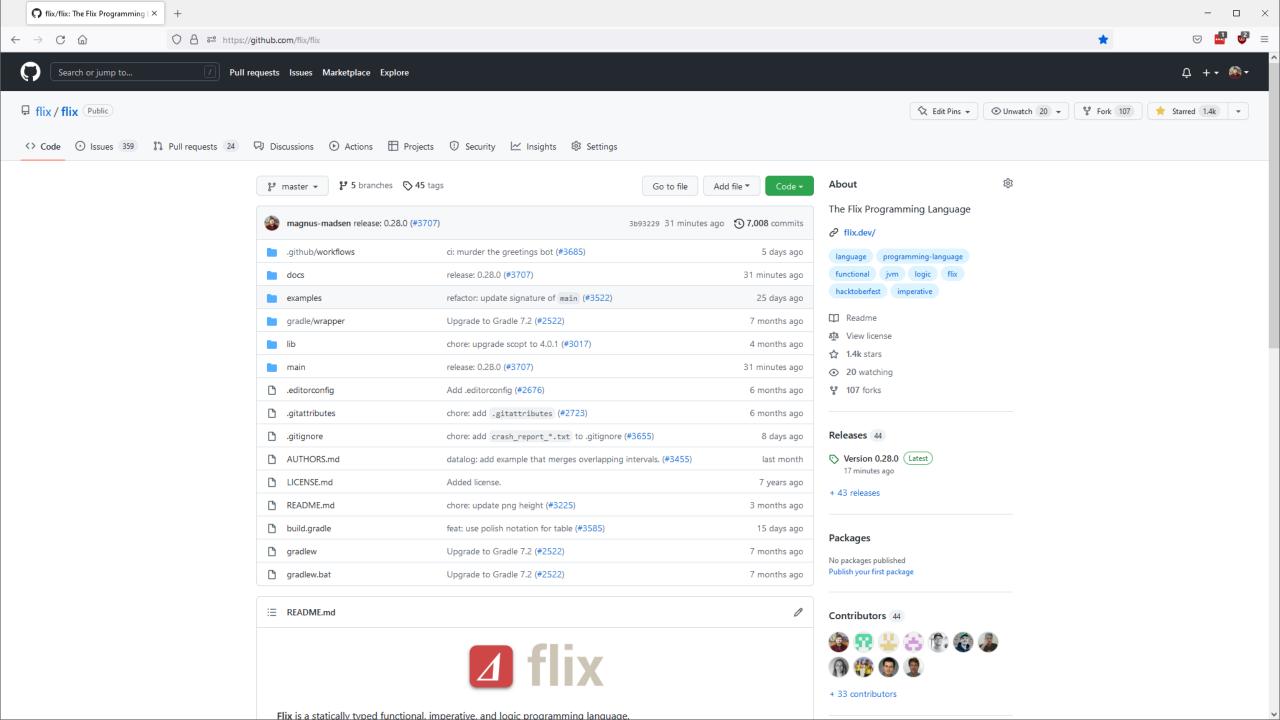
```
def isConnected(s: Set[(Int32, Int32)], src: Int32, dst: Int32): Bool =
    let rules = #{
        Path(x, y) :- Edge(x, y).
        Path(x, z) :- Path(x, y), Edge(y, z).
    };
    let edges = project s into Edge;
    let paths = query edges, rules select true from Path(src, dst);
    not (paths |> Array.isEmpty)

def main(): Unit & Impure =
    let s = Set#{(1, 2), (2, 3), (3, 4), (4, 5)};
    let src = 1;
    let dst = 5;
    if (isConnected(s, src, dst)) {
        println("Found a path between ${src} and ${dst}!")
    } else {
        println("Did not find a path between ${src} and ${dst}!")
}
```

The isConnected function behaves like any other function: We can call it with a set of edges (Int32-pairs), an Int32 source node, and an Int32 destination node. What is interesting about isConnected is that its implementation uses a small Datalog program to solve the task at hand.

In the isConnected function, the local variable rules holds a Datalog program fragment that consists of two rules which define the Path relation. Note that the predicate symbols, Edge and Path do not have to be explicitly introduced; they are simply used. The local variable edges holds a collection of edge facts that are obtained by taking all the tuples in the set sand turning them into Edge facts. Next, the local variable pathsholds the result of computing the fixpoint of the facts and rules (edges and rules) and selecting the Boolean true if there is a Path(src, dst) fact. Note that here srcand dst are the lexically-bound function parameters. Thus, pathsis either an empty array (no paths were found) or a one-element array (a path was found), and we simply return this fact.







Wrapping Up

Selection of Research Papers

From Datalog to Flix: A Declarative Language for Fixed Points on Lattices [PLDI '16]

Magnus Madsen, Ming-Ho Yee, Ondřej Lhoták

Fixpoints for the Masses: Programming with First-Class Datalog Constraints [OOPSLA '20]

Magnus Madsen, Ondřej Lhoták

Polymorphic Types and Effects with Boolean Unification [OOPSLA '20]

Magnus Madsen, Jaco van de Pol

Safe and Sound Program Analysis with Flix [ISSTA '18]

Magnus Madsen, Ondřej Lhoták

Implicit Parameters for Logic Programming [PPDP '18]

Magnus Madsen, Ondřej Lhoták

Relational Nullable Types with Boolean Unification [OOPSLA '21]

Magnus Madsen, Jaco van de Pol

You can find all papers at: https://flix.dev/research/

Additional Resources

The Official Flix Website: https://flix.dev/

The Programming Flix Book: https://doc.flix.dev/

API Documentation: https://api.flix.dev/

GitHub: https://github.com/flix/flix

InfoQ Article: https://tinyurl.com/infoq-flix

Happy Path Podcast: https://tinyurl.com/happypath-flix

Summary (1/2)

Flix is a new functional, imperative, and logic programming language.

Flix aims to offer a unique combination of features that no other existing language offers.

- algebraic data types and pattern matching
- tuples and extensible records
- parametric polymorphism
- higher-kinded types and type classes
- a polymorphic effect system
- purity reflection
- local region-based mutation

channel and process-based concurrency

first-class Datalog program values

Hindley-Milner style type inference

full tail call elimination

an extensive standard library

Visual Studio Code support

... and more

Summary (2/2)

Flix is a new functional, imperative, and logic programming language.

Flix is ready for use! Try it out!

- Flix has a fully-featured Visual Studio Code extension.
- Flix has a website, online documentation, and a playground.
- The Flix Standard Library is extensive.

Thank You!

Flix is open source, freely available, and ready for use:

https://flix.dev/