Rebuilding Constructive Type Theory with Cedille



Aaron Stump Computer Science The University of Iowa Iowa City, Iowa

Rediscovering Constructive Type Theory with Cedille



Aaron Stump Computer Science The University of Iowa Iowa City, Iowa In this talk, let us

Seek out green type theory, by going back to a simpler time, without data,

only λ .

Plan for the talk



Cedille, motivation and architecture

merge Histomorphic mergesort



Current and future directions



The mission: seek out this greener type theory,

passing through perilous, unexplored territory

following the course of

The mission: seek out this greener type theory,

passing through perilous, unexplored territory

following the course of the Missouri River...?

The object of your mission is to explore the Missouri river, & such principal stream of it, as, by it's course & communication with the water of the Pacific ocean may offer the most direct & practicable water communication across this continent, for the purposes of commerce. The object of your mission is to explore the Missouri river, & such principal stream of it, as, by it's course & communication with the water of the Pacific ocean may offer the most direct & practicable water communication across this continent, for the purposes of commerce.

Thomas Jefferson to Meriwether Lewis, June 20, 1803.

Let us form a type-theoretic Corps of Discovery!



Let us form a type-theoretic Corps of Discovery!



For a tour of the recent history of type theory





BRITISH

Pittsburgh, Pennsylvania

Great Falls

New Tour

Clabert P



TERRITORY

Charlottese

OREGON BRITISH

Pittsburgh, Pennsylvania

Classer of

Great Falls

Reynolds, Girard invent System F (1970s)

Nese Tour

Impredicative polymorphism $\forall X. F$

Beyond (current) ordinal analysis: powerful!



TERRITORY



BRITISH

Great Falls

New Tour

St. Louis, Missouri

Non Con



TERRITORY

Charlotter

OREGON BRITISH

Great Falls

St. Louis, Missouri

For Cr

Coquand, Huet: Calculus of Constructions (1988)

Nese Tour

Add dependent types $\Pi x : A. B$

No induction [Geuvers 2001]



TERRITORY



OREGON BRITISH

Council Bluffs, Iowa

Sor C

Luo: Extended Calculus of Constructions (1990) Add predicative hierarchy Prop, $Type_j$, $j \in \mathbb{N}$ Extend impredicativity $\Pi x : Type_i$. P : Prop



TERRITORY



Great Falls, Montana

OREGON

Werner: Calculus of Inductive Constructions [1994]

Add primitive inductive types

(No predicative hierarchy)

Finally ready for formalizing Math/CS!

BRITISH



PRITORY



Wait a second!



Wait a second!

Coq is fantastic, but...

We locked a fixed notion of inductive types

into our core theory.

Wait a second!

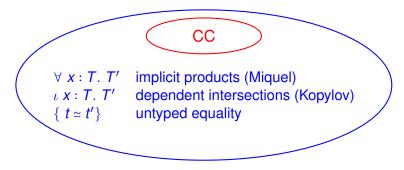
Coq is fantastic, but...

We locked a fixed notion of inductive types

into our core theory.

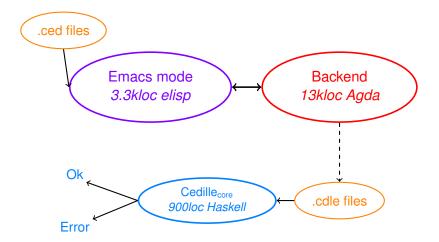
Should we maybe try the Platte through Nebraska?

Introducing Cedille



- > Small theory, formal syntax and semantics
- Core checker implemented in < 1000loc Haskell</p>
- b Logically sound
- > Turing complete(!)
- Since Cedille 1.1, datatype notations, elaborated to
- ▷ Inductive, efficient lambda-encodings

Architecture of Cedille

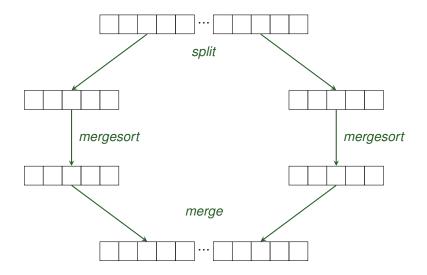


Demo

- ▷ Deriving induction
- ▷ Casts and recursive types

Histomorphic mergesort

Classic mergesort



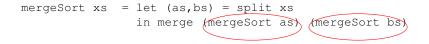
Classic mergesort (in Haskell)

From rosettacode.org

Classic mergesort is not simple in Type Theory

> Splitting and merging are structurally recursive.

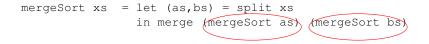
Classic mergesort is not simple in Type Theory



> Splitting and merging are structurally recursive.

▷ mergeSort itself is not.

Classic mergesort is not simple in Type Theory



 \triangleright Splitting and merging are structurally recursive.

▷ mergeSort itself is not.

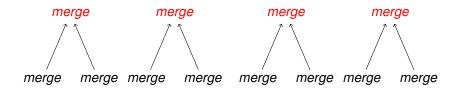
To rectify, various techniques can be applied:

- ▷ well-founded recursion
- ▷ sized types [Copello et al. 2014]
- inductive domains (cf. great survey paper "Partiality and Recursion in ITPs", [Bove et al. 2016])

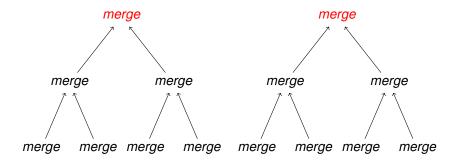
Bottom-up mergesort: a balanced tree of merges

merge merge merge merge merge merge merge

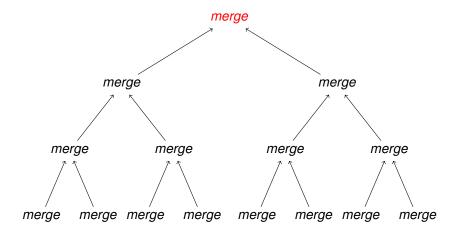
Bottom-up mergesort: a balanced tree of merges



Bottom-up mergesort: a balanced tree of merges



Bottom-up mergesort: a balanced tree of merges



Bottom-up mergesort also bad for TT

```
mergeSortBottomUp list = mergeAll (map (x \rightarrow [x]) list)
```

```
mergeAll [sorted] = sorted
mergeAll sorteds = mergeAll (mergePairs sorteds)
```

```
mergePairs (s1 : s2 : ss) = merge s1 s2 : mergePairs ss
mergePairs sorteds = sorteds
```

From rosettacode.org

Bottom-up mergesort also bad for TT

```
mergeSortBottomUp list = mergeAll (map (\x -> [x]) list)
mergeAll [sorted] = sorted
mergeAll sorteds = mergeAll (mergePairs sorteds)
mergePairs (s1 : s2 : ss) = merge s1 s2 : mergePairs ss
mergePairs sorteds = sorteds
```

From rosettacode.org

> mergeAll not structurally recursive

Bottom-up mergesort also bad for TT

```
mergeSortBottomUp list = mergeAll (map (\x -> [x]) list)
mergeAll [sorted] = sorted
mergeAll sorteds = mergeAll (mergePairs sorteds)
mergePairs (s1 : s2 : ss) = merge s1 s2 : mergePairs ss
mergePairs sorteds = sorteds
```

From rosettacode.org

> mergeAll not structurally recursive

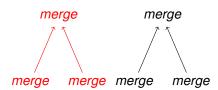
b though mergePairs decreases size of list (if non-nil)

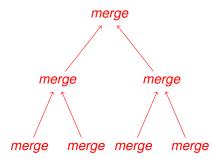
merge

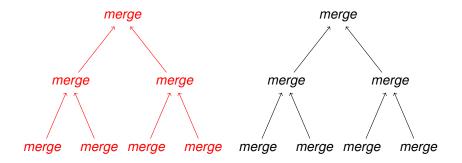
merge merge

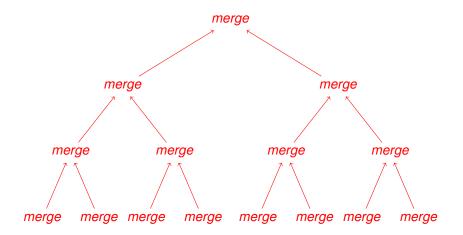
merge

merge merge





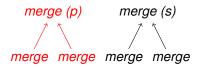






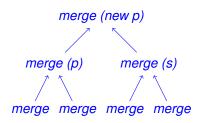
For increasing *k* starting from 0:

 \triangleright Let *p* be prefix (length 2^{*k*}) already sorted



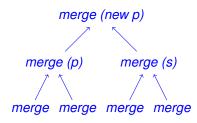
For increasing *k* starting from 0:

- \triangleright Let *p* be prefix (length 2^{*k*}) already sorted
- \triangleright Let *s* be bottom-up mergesort of next 2^{*k*} elements



For increasing *k* starting from 0:

- \triangleright Let *p* be prefix (length 2^{*k*}) already sorted
- \triangleright Let *s* be bottom-up mergesort of next 2^{*k*} elements
- \triangleright Update *p* to merge of *p* and *s*



For increasing *k* starting from 0:

- \triangleright Let *p* be prefix (length 2^{*k*}) already sorted
- \triangleright Let *s* be bottom-up mergesort of next 2^{*k*} elements

...

- \triangleright Update *p* to merge of *p* and *s*
- ▷ Stop when list is empty

```
data Nat = Zero | Succ Nat
takePow2 :: Nat -> Int -> [Int] -> ([Int],[Int])
takePow2 = ...
prefixMergeSort :: [Int] -> [Int]
prefixMergeSort [] = []
prefixMergeSort (h : t) = loop t [h] Zero
where loop [] p _ = p
loop (h : t) p n =
let (s,t') = takePow2 n h t in
loop t' (merge p s) (Succ n)
```

```
data Nat = Zero | Succ Nat
takePow2 :: Nat -> Int -> [Int] -> ([Int],[Int])
takePow2 = ...
prefixMergeSort :: [Int] -> [Int]
prefixMergeSort [] = []
prefixMergeSort (h : t) = loop t [h] Zero
where loop [] p _ = p
loop (h : t) p n =
let (s,t') = takePow2 n h t in
loop t' (merge p s) (Succ n)
```

▷ takePow2 structurally recursive on the Nat

data Nat = Zero | Succ Nat takePow2 :: Nat -> Int -> [Int] -> ([Int], [Int]) takePow2 = ... prefixMergeSort :: [Int] -> [Int] prefixMergeSort [] = [] prefixMergeSort (h : t) = loop t [h] Zero where loop [] p _ = p loop (h : t) p n = let (s,t') = takePow2 n h t in loop t' (merge p s) (Succ n)

D takePow2 structurally recursive on the Nat

D takePow2 takes head and tail, to assure that

data Nat = Zero | Succ Nat takePow2 :: Nat -> Int -> [Int] -> ([Int], [Int]) takePow2 = ... prefixMergeSort :: [Int] -> [Int] prefixMergeSort [] = [] prefixMergeSort (h : t) = loop t [h] Zero where loop [] p _ = p loop (h : t) p n = let (s,t') = takePow2 n h t in loop t' (merge p s) (Succ n)

> takePow2 structurally recursive on the Nat

> takePow2 takes head and tail, to assure that

▷t' is a sublist of t

data Nat = Zero | Succ Nat takePow2 :: Nat -> Int -> [Int] -> ([Int], [Int]) takePow2 = ... prefixMergeSort :: [Int] -> [Int] prefixMergeSort [] = [] prefixMergeSort (h : t) = loop t [h] Zero where loop [] p _ = p loop (h : t) p n = let (s.t') = takePow2 n h t in loop t' (merge p s) (Succ n)

▷ takePow2 structurally recursive on the Nat

> takePow2 takes head and tail, to assure that

▷ t' is a sublist of t

 \triangleright So recursive call to <code>loop</code> is structurally decreasing (h : t \rightarrow t')

How do we code this in Cedille?

- Cedille implements histomorphic recursion
- > Pattern-matching recursions on inductive data D provide
 - An abstract type A
 - A function to use for recursive calls on type A
 - Evidence that A is D-like
 - can be decomposed like D, and
 - cast to D
- ▷ Datatype declarations add predicate for *D*-like, some helpers

The List datatype

```
data List (A: *): * =

| nil: List

| cons: A \rightarrow List \rightarrow List.
```

This declaration introduces:

Is/List : $\Pi A : \star . \star \to \star$ predicate for list-like

The List datatype

```
data List (A: *): * =

| nil: List

| cons: A \rightarrow List \rightarrow List.
```

This declaration introduces:

Is/List: Π A : * . * \rightarrow *predicate for list-likeis/List: \forall A : * . Is/List \cdot A \cdot (List \cdot A)evidence that List is list-like

The List datatype

```
data List (A: *): * =

| nil: List

| cons: A \rightarrow List \rightarrow List.
```

This declaration introduces:

The loop helper function for prefix mergesort <u>Haskell:</u>

```
loop [] p _ = p
loop (h : t) p n =
    let (s,t') = takePow2 n h t in
        loop t' (merge p s) (Succ n)
```

Cedille:

$$\begin{array}{l} \mu \mbox{ loop . t} \\ \mbox{@ } (\lambda _ : \mbox{List } \cdot \mbox{Nat } \rightarrow \mbox{Nat } \rightarrow \mbox{List } \cdot \mbox{Nat)} \\ \mbox{\{ nil } \rightarrow \mbox{λ p } . \mbox{λ } _ . \mbox{p} \\ \mbox{| cons h t } \rightarrow \mbox{λ p } . \mbox{λ n } . \\ \mbox{μ' takePow2 - isType/loop n h t} \\ \mbox{{\{ pair s t' } \rightarrow \mbox{ loop t' (merge p s) (succ n) } \}} \end{array}$$

- $\triangleright \mu$ introduces pattern-matching recursion
 - always over some given data, here a list t
- $\triangleright \mu'$ is a simple pattern-match

$$\begin{array}{l} \mu \text{ loop . t} \\ @ (\lambda _ : \text{List } \cdot \text{Nat } . \text{List } \cdot \text{Nat } \to \text{Nat } \to \text{List } \cdot \text{Nat}) \\ \{ \text{ nil } \to \lambda \text{ p } . \lambda _ . \text{ p} \\ | \text{ cons h } t \to \lambda \text{ p } . \lambda \text{ n } . \\ \mu' \text{ takePow2 } -\text{isType/loop n h } t \\ \{ \text{ pair s } t' \to \text{ loop } t' \text{ (merge p s) (succ n) } \} \end{array}$$

The following are available in the body of cons-clause:

Type/loop : *

abstract type for subdata

$$\begin{array}{l} \mu \text{ loop . t} \\ @ (\lambda _ : \text{List } \cdot \text{Nat } . \text{List } \cdot \text{Nat } \to \text{Nat } \to \text{List } \cdot \text{Nat}) \\ \{ \text{ nil } \to \lambda \text{ p } . \lambda _ . \text{ p} \\ | \text{ cons h } t \to \lambda \text{ p } . \lambda \text{ n } . \\ \mu' \text{ takePow2 } -\text{isType/loop n h } t \\ \{ \text{ pair s } t' \to \text{ loop } t' \text{ (merge p s) (succ n) } \} \end{array}$$

The following are available in the body of cons-clause:

Type/loop	:	*	abstract type for subdata
t	:	Type/loop	subdata (tail of list), at abstract type

$$\begin{array}{l} \mu \ \text{loop .t} \\ \textcircled{0}{0} (\lambda _ : \text{List } \cdot \text{Nat } . \text{List } \cdot \text{Nat } \rightarrow \text{Nat } \rightarrow \text{List } \cdot \text{Nat}) \\ \textcircled{0}{0} (\lambda _ : \text{List } \cdot \text{Nat } . \text{List } \cdot \text{Nat } \rightarrow \text{List } \cdot \text{Nat}) \\ \textcircled{0}{0} (\lambda _ : \text{List } \cdot \text{Nat } . \text{List } \cdot \text{Nat } \rightarrow \text{List } \cdot \text{Nat}) \\ \fbox{0}{0} (\lambda _ : \text{List } \cdot \text{Nat } . \text{List } \cdot \text{Nat } \rightarrow \text{List } \rightarrow \text{List } \cdot \text{Nat } \rightarrow \text{List } \rightarrow \text{List$$

The following are available in the body of cons-clause:

Type/loop	:	*	abstract type for subdata
t	:	Type/loop	subdata (tail of list), at abstract type
loop	:	Type/loop \rightarrow	List \cdot Nat \rightarrow Nat \rightarrow List \cdot Nat
			for recursive calls

$$\begin{array}{l} \mu \ \text{loop .t} \\ \textcircled{0}{0} (\lambda _ : \text{List } \cdot \text{Nat } . \text{List } \cdot \text{Nat } \rightarrow \text{Nat } \rightarrow \text{List } \cdot \text{Nat}) \\ \textcircled{0}{0} (\lambda _ : \text{List } \cdot \text{Nat } . \text{List } \cdot \text{Nat } \rightarrow \text{List } \cdot \text{Nat}) \\ \textcircled{0}{0} (\lambda _ : \text{List } \cdot \text{Nat } . \text{List } \cdot \text{Nat } \rightarrow \text{List } \cdot \text{Nat}) \\ \fbox{0}{0} (\lambda _ : \text{List } \cdot \text{Nat } . \text{List } \cdot \text{Nat } \rightarrow \text{List } \rightarrow \text{List } \cdot \text{Nat } \rightarrow \text{List } \rightarrow \text{List$$

The following are available in the body of cons-clause:

Type/loop	:	* abstract type for subdata
t	:	Type/loopsubdata (tail of list), at abstract type
loop	:	Type/loop \rightarrow List \cdot Nat \rightarrow Nat \rightarrow List \cdot Nat
		for recursive calls
isType/loop	:	Is/List ·Nat ·Type/loop
		avidance that abatract type is list like

evidence that abstract type is list-like

takePow2, using Is/List

```
takePow2 : ∀ T : * . Is/List • Nat • T ⇒
                Nat \rightarrow Nat \rightarrow T \rightarrow Pair \cdot (List \cdot Nat) \cdot T =
  \Lambda T . \Lambda mT . \lambda n .
       \mu takePow2 . n
            { zero \rightarrow \lambda a . \lambda ] . pair (singleton a) ]
            | succ n \rightarrow \lambda a . \lambda l .
                [p = takePow2 n a 1] -
                 μ′ p
                 { pair t1 1 \rightarrow
                    \mu' < mT > 1
                     \{ nil \rightarrow p \}
                      | cons a | \rightarrow
                         \mu' (takePow2 n a l)
                          { pair t2 1 \rightarrow
                            pair (merge t1 t2) 1 }}}.
```

takePow2, using Is/List

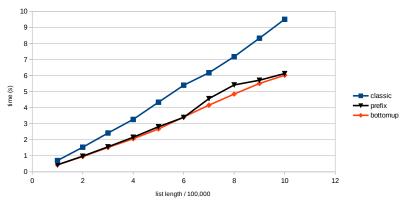
```
takePow2 : ∀ T : * . Is/List • Nat • T ⇒
                Nat \rightarrow Nat \rightarrow T \rightarrow Pair \cdot (List \cdot Nat) \cdot T =
  \Lambda T . \Lambda mT . \lambda n .
       \mu takePow2 . n
            { zero \rightarrow \lambda a . \lambda l . pair (singleton a) l
            | succ n \rightarrow \lambda a . \lambda l .
                [p = takePow2 n a 1] -
                 μ' p
                  { pair t1 l \rightarrow
                    (1 <mT>)1
                      \{ nil \rightarrow p \}
                      | cons a | \rightarrow
                         \mu' (takePow2 n a l)
                           { pair t2 1 \rightarrow
                             pair (merge t1 t2) 1 }}}.
```

- $\triangleright \mu'$ accepts evidence that T is list-like!
- ▷ So you can use pattern-matching on *D*-like types

Types for histomorphic mergesort

- \triangleright Use Is/List and μ' to define takePow2 outside of msort
- > msort can recurse on values of type Type/List returned by takePow2
- > Nothing else needed to convince the (implicit) termination checker!
- > More flexible than nested recursions/lexicographic recursions

Comparing mergesort variants in Haskell



For each list length,

- Randomly generate 10 lists of that length
- > Run each sorting algorithm (as a separate process)
- List elements range from min to max Int



Schematic Cedille (Cedille 1.2)

When you run out of power...

Because you must at some point, if sound [Gödel 1931]

- ▷ Predicative hierarchy?
 - Complexities with level expressions
 - Temptation to keep going (universe-polymorphism!)
- > A time-honored alternative:

When you run out of power...

Because you must at some point, if sound [Gödel 1931]

- Predicative hierarchy?
 - Complexities with level expressions
 - Temptation to keep going (universe-polymorphism!)
- > A time-honored alternative: go schematic!

What does schematic mean?

Allowing parametrized definitions not expressible in the language.

In type theory: definitions exceeding the allowed products.

What does schematic mean?

Allowing parametrized definitions not expressible in the language. In type theory: definitions exceeding the allowed products.

A simple example in current Cedille: parametrized kind definitions.

 $\kappa (I : \star) = I \rightarrow \star.$

- \triangleright In Cedille, λ I : \star . I \rightarrow \star is not typable (no $\star \rightarrow \Box$)
- $\triangleright~$ All uses of κ must include an argument for ${\tt I}$
- > So such definitions work like typed macros
- ▷ But we want to go beyond this...

Telescope-generic developments

- > We are working on parametrization by telescopes
 - A telescope is a dependent sequence of declarations, e.g.

(A : *) (a : A)

- Allow definitions and modules to be parametric in $\gamma~:~{\tt tel}$
- Form products over such γ , with λ -abstractions and applications
- ⊳ Benefits:
 - More generic developments (avoid duplication), especially
 - Lambda-encodings with different parameter/index lists
 - Eliminate complex code in Cedille for translating datatypes,
 - Replacing with telescope-generic Cedille

Schematic RecType

For recursive indexed types:

Instantiate γ by (n : Nat) for Nat-indexed recursive types (e.g.)

Native disjunctions and existentials (Cedille 1.3)

We can already encode existential types

```
module WeakSigma(A : *) (B : A \rightarrow *).
Exists : * = \forall X : * . (\forall a : A . B a \rightarrow X) \rightarrow X .
witness : \forall a : A . B a \rightarrow wSigma =
\Lambda a . \lambda b . \Lambda X . \lambda c . c -a b .
```

- \triangleright But witness –a b normalizes to λ c . c b
- > For some situations, you really want just b
- ▷ Also, we are considering *n*-ary disjunctions
 - Introductions inj_k t
 - As opposed to inj₂ (··· inj₂ (inj₁t)) for binary
- Long bad history in proof theory...

The unpleasant eliminations

$$\frac{\Gamma \vdash t : \exists x : T.F \qquad \Gamma, x : T, u : F \vdash t' : C}{\Gamma \vdash unpack \ t \ as \ (x.y.t') : C}$$

$$\frac{\Gamma \vdash t : F_1 \lor F_2 \qquad \Gamma, u : F_1 \vdash t_1 : C \qquad \Gamma, v : F_2 \vdash t_2 : C}{\Gamma \vdash case \ t \ of \ (u.t_1, v.t_2) : C}$$

- ▷ Host of well-known issues (commuting conversions, etc.)
- ▷ Like a cut of sequent calculus

The unpleasant eliminations

$$\frac{\Gamma \vdash t : \exists x : T.F \qquad \Gamma, x : T, u : F \vdash t' : C}{\Gamma \vdash unpack \ t \ as \ (x.y.t') : C}$$

$$\frac{\Gamma \vdash t : F_1 \lor F_2 \qquad \Gamma, u : F_1 \vdash t_1 : C \qquad \Gamma, v : F_2 \vdash t_2 : C}{\Gamma \vdash case \ t \ of \ (u.t_1, v.t_2) : C}$$

- ▷ Host of well-known issues (commuting conversions, etc.)
- ▷ Like a cut of sequent calculus ...

The unpleasant eliminations

$$\frac{\Gamma \vdash t : \exists x : T.F \qquad \Gamma, x : T, u : F \vdash t' : C}{\Gamma \vdash unpack \ t \ as \ (x.y.t') : C}$$

$$\frac{\Gamma \vdash t : F_1 \lor F_2 \qquad \Gamma, u : F_1 \vdash t_1 : C \qquad \Gamma, v : F_2 \vdash t_2 : C}{\Gamma \vdash case \ t \ of \ (u.t_1, v.t_2) : C}$$

- ▷ Host of well-known issues (commuting conversions, etc.)
- ▷ Like a cut of sequent calculus ... !

Sequent calculus and natural deduction

> Natural deduction as emerging from sequent calculus

- (Historically the reverse.)
- Left-rules of seq. calc. give rise (somehow!) to eliminations
- We pivot seq. calc. proofs so left-rules go to the top
- Eliminations followed by introductions (normal proofs)
- $\,\triangleright\,$ For good connectives, this works ($\rightarrow,\,\wedge,\,\intercal,\,\bot,\,\forall)$
- \triangleright For \lor and \exists , we get stuck

Let us assign natural-deduction terms to sequent proofs

Right-rules get assigned introduction forms

Left-rules get elimination forms... but where? how?

Let us assign natural-deduction terms to sequent proofs

Co-projections

$$\frac{\Gamma, t.(1): T_1 \vdash t_1: C \qquad \Gamma, t.(2): T_2 \vdash t_2: C}{\Gamma, t: T_1 \lor T_2 \vdash t_1 || t_2: C} \lor L$$

Eliminations t.(1) and t.(2), introductions $inj_1 t$, $inj_2 t$

The $(\lor L)$ rule ensures that one branch of $t_1 || t_2$ will succeed

▷ Failure is reducing $(inj_1 t).(2)$ or $(inj_2 t).(1)$

Not sure yet how to formulate natural-deduction typing for these

Switch to Sequent Calculus (Cedille 2.0)

Cuts and control!

- > Cuts give rise to control operators classically
- Good reasons for avoid classicality for programs
 - Under Curry-Howard, $T \lor T'$ is a sum type
 - But classically, get no information splitting on $T \lor \neg T$
- > Can we support control and retain canonicity?
- > Can we also achieve perfect duality?
 - Logics like BiInt have duality and control, but
 - Not canonicity!
 - Can prove $F \lor (\top \prec F)$
- ▷ Modifying a system of Wansing's, we have a candidate...

Cuts and control!

- > Cuts give rise to control operators classically
- Good reasons for avoid classicality for programs
 - Under Curry-Howard, $T \lor T'$ is a sum type
 - But classically, get no information splitting on $T \lor \neg T$
- > Can we support control and retain canonicity?
- > Can we also achieve perfect duality?
 - Logics like BiInt have duality and control, but
 - Not canonicity!
 - Can prove $F \lor (\top \prec F)$
- Modifying a system of Wansing's, we have a candidate... to be discussed another time!

Conclusion: our route

Following the Missouri River (CC),

We took the North Platte from Omaha (Curry-style *CC* + *ι*, ≃, ∀)
 Heading south (hotter? histomorphic recursion)
 to the Colorado River (getting crazy)
 towards the Pacific (ultimate Type Theory)

Right now we are maybe passing through...



Acknowledgments

- ▷ Current Cedille dev team:
 - Postdoc: Stephan Spahn
 - Doctoral: Andrew Marmaduke, Chris Jenkins, Tony Cantor
 - High schooler: Colin McDonald
- Group alums: Denis Firsov, Larry Diehl, Ernesto Copello, Richard Blair, Ananda Guneratne, Chad Reynolds, Matthew Heimerdinger
- Funders: NSF 1524519, DoD FA9550-16-1-0082

