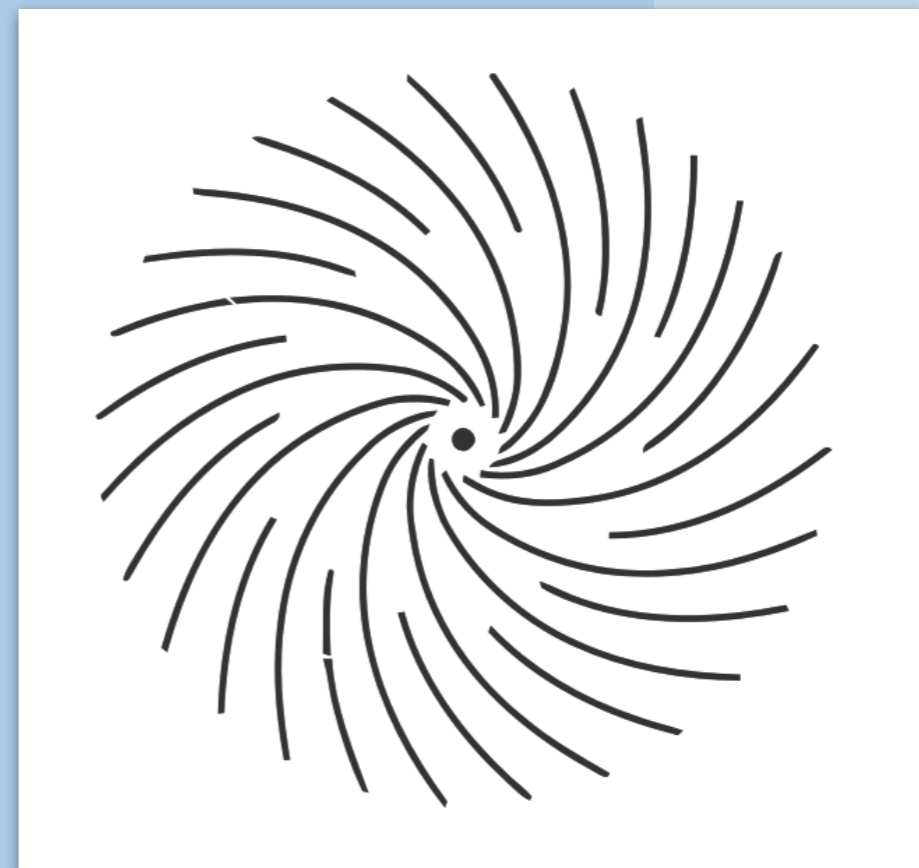


6. Fixed Points

Oscar Nierstrasz



Roadmap



- > Representing Numbers
- > Recursion and the Fixed-Point Combinator
- > The typed lambda calculus
- > The polymorphic lambda calculus
- > Other calculi

References

- > Paul Hudak, “*Conception, Evolution, and Application of Functional Programming Languages*,” ACM Computing Surveys 21/3, Sept. 1989, pp 359-411.

Conception, Evolution, and Application of Functional Programming Languages

<http://scgresources.unibe.ch/Literature/PL/Huda89a-p359-hudak.pdf>

Roadmap



- > **Representing Numbers**
- > Recursion and the Fixed-Point Combinator
- > The typed lambda calculus
- > The polymorphic lambda calculus
- > Other calculi

Recall these encodings ...

True $\equiv \lambda x y . x$

False $\equiv \lambda x y . y$

pair $\equiv (\lambda x y z . z x y)$

(x, y) \equiv pair x y

first $\equiv (\lambda p . p \text{ True })$

second $\equiv (\lambda p . p \text{ False })$

Representing Numbers

There is a “standard encoding” of natural numbers into the lambda calculus:

Define:

$$0 \equiv (\lambda x . x)$$
$$\text{succ} \equiv (\lambda n . (\text{False}, n))$$

then:

1	\equiv	succ 0	\rightarrow	(False, 0)
2	\equiv	succ 1	\rightarrow	(False, 1)
3	\equiv	succ 2	\rightarrow	(False, 2)
4	\equiv	succ 3	\rightarrow	(False, 3)

Note how all positive integers are encoded as a pair (b, n) , where b is a Boolean indicating whether the number is zero or not (i.e., always *False* for positive integers), and n represents the number's predecessor. In other words, the number $n+1$ is represented by $(False, n)$, which is computed by the function application $succ\ n$.

The exception is the integer value 0, which is represented by the identify function.

Now the question is, how do we make sense of this representation of 0?

Working with numbers

We can define simple functions to work with our numbers.

Consider:

iszero \equiv first

pred \equiv second

then:

iszero 1 = first (False, 0) \rightarrow False

iszero 0 = $(\lambda p . p \text{ True}) (\lambda x . x)$ \rightarrow True

pred 1 = second (False, 0) \rightarrow 0

 *What happens when we apply pred 0? What does this mean?*

To work with our numbers, we need to do at least two further things: check if a number is zero or not, and compute its predecessor.

If we define *iszero* as the function *first* (seen earlier) then we see that it works correctly both for positive integers and the value 0, returning *False* for the former and *True* for the latter.

(Ok, it's a trick, but it's a cool trick!)

Similarly, if we define *pred* as *second*, it correctly returns the predecessor of all positive integers.

What, however, is the meaning of *pred 0* ?

Roadmap



- > Representing Numbers
- > **Recursion and the Fixed-Point Combinator**
- > The typed lambda calculus
- > The polymorphic lambda calculus
- > Other calculi

Recursion

Suppose we want to define *arithmetic operations* on our lambda-encoded numbers.

In Haskell we can program:

```
plus n m
  | n == 0      = m
  | otherwise   = plus (n-1) (m+1)
```

so we might try to “define”:

$$\text{plus} \equiv \lambda n m . \text{iszero } n \ m \ (\text{plus } (\text{pred } n) \ (\text{succ } m))$$

Unfortunately this is *not a definition*, since we are trying to *use plus before it is defined*. I.e, plus is *free* in the “definition”!

The problem is that \equiv is not part of the syntax of the lambda calculus: we have no way to define a “global function”, but can only work with anonymous lambdas. The only way we can bind a name to a function is using a lambda itself. For example, if we want to evaluate *not True*, then we need to first bind the names *True* and *False* using lambdas, then *not* using *True* and *False*, and finally evaluate *not True*:

$(\lambda \text{ True False .$

$(\lambda \text{ not . not True)$

$(\lambda \text{ b.b False True)$

$) (\lambda \text{ x y.x}) (\lambda \text{ x y.y})$

With *plus* this trick will not work, since we would need to bind *plus* before we could use it, which is impossible!

Recursive functions as fixed points

We can obtain a closed expression by *abstracting over plus*:

$$\text{rplus} \equiv \lambda \text{ plus } n \ m . \text{iszero } n \\ \quad \quad \quad m \\ \quad \quad \quad (\text{plus } (\text{pred } n) (\text{succ } m))$$

rplus takes as its *argument* the actual plus function to use and returns as its result a definition of that function in terms of itself. In other words, if **fplus** is the function we want, then:

$$\text{rplus fplus} \leftrightarrow \text{fplus}$$

I.e., we are searching for a *fixed point* of rplus ...

The trick here is to *abstract away from plus*, i.e., to consider it as a *parameter* to the function *plus* we are defining. Although this seems circular, it is not. We are saying how to *define plus*, if we are given an *implementation* of *plus*. These are two different things.

Now the question is how we can turn this definition into the implementation that we need!

The missing implementation is a “fixed point” of the abstraction $rplus$: if we supply it as an argument, we get the same value back as a result.

Fixed Points

A fixed point of a function f is a value p such that $f\ p = p$.

Examples:

`fact 1 = 1`

`fact 2 = 2`

`fib 0 = 0`

`fib 1 = 1`

Fixed points are not always “well-behaved”:

`succ n = n + 1`

 *What is a fixed point of succ?*

Fixed Point Theorem

Theorem:

Every lambda expression e has a fixed point p such that $(e\ p) \leftrightarrow p$.

Proof:

Let: $Y \equiv \lambda f . (\lambda x . f (x\ x)) (\lambda x . f (x\ x))$

Now consider:

$$\begin{aligned} p \equiv Y\ e &\rightarrow (\lambda x . e\ (x\ x)) (\lambda x . e\ (x\ x)) \\ &\rightarrow e\ ((\lambda x . e\ (x\ x)) (\lambda x . e\ (x\ x))) \\ &= e\ p \end{aligned}$$

So, the “magical Y combinator” can always be used to find a fixed point of an *arbitrary* lambda expression.

$$\forall e: Y\ e \leftrightarrow e\ (Y\ e)$$

The “magical” Y combinator will take any lambda expression e , and yield a value $Y e$, which, when passed to e as an argument, returns a result that is *equivalent* (i.e., convertible with reductions and conversions) to $Y e$. $Y e$ is therefore a *fixpoint* (AKA “fixed point”) of e .

Y is “magical” in the sense that it is not immediately obvious why it should work, but it is not too hard to see why ...

How does Y work?

Recall the non-terminating expression

$$\Omega = (\lambda x . x x) (\lambda x . x x)$$

Ω loops endlessly without doing any productive work.

Note that $(x x)$ represents the body of the “loop”.

We simply define Y to take an *extra parameter* f , and *put it into the loop*, passing it the body as an argument:

$$Y \equiv \lambda f . (\lambda x . f (x x)) (\lambda x . f (x x))$$

So Y just inserts some productive work into the body of Ω

Note how Ω just loops around endlessly without doing anything useful. The expression $(x\ x)$ represents this loop.

Recall that $Y\ e$ is the fixpoint we seek. We define Y just like Ω , but we add an extra parameter f (which will be bound to e), and we pass it the result of the loop, i.e., $(x\ x)$. Now it is up to e to decide when to terminate the loop!

If e is defined recursively, it will loop as often as it needs to, and then end with the result it is supposed to compute. By the trick of self-application in $(x\ x)$, it can get as many copies of the fixpoint as it needs, but it no longer has to loop endlessly.

Using the Y Combinator

Consider:


$f \equiv \lambda x. \text{True}$

then:

$Y f \rightarrow f (Y f)$ *by FP theorem*
 $= (\lambda x. \text{True}) (Y f)$
 $\rightarrow \text{True}$

Consider:

$Y \text{succ} \rightarrow \text{succ} (Y \text{succ})$ *by FP theorem*
 $\rightarrow (\text{False}, (Y \text{succ}))$

 *What are succ and pred of (False, (Y succ))? What does this represent?*

Recall that the fixpoint theorem states that

$$Y\ e \rightarrow e\ (Y\ e)$$

We simply apply this rewriting and evaluate the resulting expression. Note that for any constant function $(\lambda\ x.\ k)$ (where k is a constant):

$$Y\ (\lambda\ x.\ k) \rightarrow (\lambda\ x.\ k)\ (Y\ (\lambda\ x.\ k)) \rightarrow k$$

In other words, the fixpoint of a constant function is that constant, which is as we expect.

Recursive Functions are Fixed Points

We seek a fixed point of:

$$\text{rplus} \equiv \lambda \text{ plus } n \ m . \text{iszero } n \ m \ (\text{plus } (\text{pred } n) \ (\text{succ } m))$$

By the Fixed Point Theorem, we simply define:

$$\text{plus} \equiv Y \ \text{rplus}$$

Since this guarantees that:

$$\text{rplus } \text{plus} \leftrightarrow \text{plus}$$

as desired!

Note that the “new” `plus` that we are defining globally is not the same as the locally bound name `plus` within `rpplus`, but when we apply `rpplus plus`, the globally defined one gets bound to the local one.

Unfolding Recursive Lambda Expressions

plus 1 1 = (Y rplus) 1 1
→ rplus plus 1 1 *(NB: fp theorem)*
→ iszero 1 1 (plus (pred 1) (succ 1))
→ **False 1 (plus (pred 1) (succ 1))**
→ plus (pred 1) (succ 1)
→ rplus plus (pred 1) (succ 1)
→ iszero (pred 1) (succ 1)
 (plus (pred (pred 1)) (succ (succ 1)))
→ iszero 0 (succ 1) (...)
→ **True (succ 1) (...)**
→ succ 1
→ 2

Here we can see how, like Ω , *Y rplus* loops repeatedly. The difference is that we can do productive work at each step, and decide when to terminate the loop with the *iszero* test.

Note that we use the fixpoint theorem also in lines 5 and 6 to rewrite *plus* as *rplus plus*.

Roadmap



- > Representing Numbers
- > Recursion and the Fixed-Point Combinator
- > **The typed lambda calculus**
- > The polymorphic lambda calculus
- > Other calculi

The Typed Lambda Calculus

There are many variants of the lambda calculus.

The typed lambda calculus just *decorates terms with type annotations*:

Syntax:

$$e ::= x^\tau \mid e_1^{\tau_2 \rightarrow \tau_1} e_2^{\tau_2} \mid (\lambda x^{\tau_2}. e^{\tau_1})^{\tau_2 \rightarrow \tau_1}$$

Operational Semantics:

$$\begin{aligned} \lambda x^{\tau_2}. e^{\tau_1} &\Leftrightarrow \lambda y^{\tau_2}. [y^{\tau_2}/x^{\tau_2}] e^{\tau_1} && y^{\tau_2} \text{ not free in } e^{\tau_1} \\ (\lambda x^{\tau_2}. e_1^{\tau_1}) e_2^{\tau_2} &\Rightarrow [e_2^{\tau_2}/x^{\tau_2}] e_1^{\tau_1} \\ \lambda x^{\tau_2}. (e^{\tau_1} x^{\tau_2}) &\Rightarrow e^{\tau_1} && x^{\tau_2} \text{ not free in } e^{\tau_1} \end{aligned}$$

Example:

$$\text{True} \equiv (\lambda x^A. (\lambda y^B. x^A)^{B \rightarrow A}) A \rightarrow (B \rightarrow A)$$

This looks complicated due to the messy annotations, but the operational semantics is exactly the same as in the untyped lambda calculus. The α , β and η rules work just as before, they just keep track of additional type information, encoding for each expression and subexpression what its type is.

Note that this is a *monomorphic type system*: every value has a unique, concrete type, not a variable type. The function

$$\lambda x^A . x^A$$

Takes arguments only of type A , not any other type!

Roadmap



- > Representing Numbers
- > Recursion and the Fixed-Point Combinator
- > The typed lambda calculus
- > **The polymorphic lambda calculus**
- > Other calculi

The Polymorphic Lambda Calculus

Polymorphic functions like “map” cannot be typed in the typed lambda calculus!

Need *type variables* to capture polymorphism:

β reduction (ii):

$$(\lambda x^v . e_1^{\tau_1}) e_2^{\tau_2} \Rightarrow [\tau_2/v] [e_2^{\tau_2}/x^v] e_1^{\tau_1}$$

Example:

$$\begin{aligned} \text{True} &\equiv (\lambda x^\alpha . (\lambda y^\beta . x^\alpha)^{\beta \rightarrow \alpha})^{\alpha \rightarrow (\beta \rightarrow \alpha)} \\ \text{True}^{\alpha \rightarrow (\beta \rightarrow \alpha)} a^A b^B &\rightarrow (\lambda y^\beta . a^A)^{\beta \rightarrow A} b^B \\ &\rightarrow a^A \end{aligned}$$

NB: some superscripts have been left off for readability!

Type variables solve the problem with the monomorphic typed lambda calculus. Now we can write:

$$\lambda x^\alpha . x^\alpha$$

where α is a *type variable*. This function will accept an argument of any type A , which will then be *bound* to α . In other words, if this function accepts an argument of type A , it will return an value x of the *same* type A .

Hindley-Milner Polymorphism

Hindley-Milner polymorphism (i.e., that adopted by ML and Haskell) works by inferring the type annotations for a slightly restricted subcalculus: polymorphic functions.

If:

```
doubleLen len len' xs ys = (len xs) + (len' ys)
```

then

```
doubleLen length length "aaa" [1,2,3]
```

is ok, but if

```
doubleLen' len xs ys = (len xs) + (len ys)
```

then

```
doubleLen' length "aaa" [1,2,3]
```

is a type error since the argument `len` *cannot be assigned a unique type!*

Polymorphism and self application

Even the polymorphic lambda calculus is not powerful enough to express certain lambda terms.

Recall that both Ω and the Y combinator make use of “self application”:

$$\Omega = (\lambda x . \mathbf{x x}) (\lambda x . \mathbf{x x})$$

 *What type annotation would you assign to $(\lambda x . \mathbf{x x})$?*

Note that there is *no reasonable type* that we can assign to an expression like $x\ x$. It is therefore impossible to express Y in the typed lambda calculus, so there is no way to express recursion, and consequently, one can only express *terminating* programs using it!

For the same reason, we cannot write Y in Haskell (try it!).

Built-in recursion with letrec AKA def AKA μ

- > Most programming languages provide direct support for recursively-defined functions (avoiding the need for Y)

(def f.E) e \rightarrow **([(def f.E) / f] E) e**

(def plus. λ n m . iszero n m (plus (pred n) (succ m))) 2 3
 \rightarrow (λ n m . iszero n m ((def plus. ...) (pred n) (succ m))) 2 3
 \rightarrow (iszero 2 3 ((def plus. ...) (pred 2) (succ 3)))
 \rightarrow ((def plus. ...) (pred 2) (succ 3))
 \rightarrow ...

The “def” feature allows you to assign a name (such as f) to a lambda expression E , where f may be used recursively within E . The β reduction rule for this extended lambda calculus simply propagates the definition to the recursive uses of f within E (exactly as Y does!).

Roadmap



- > Representing Numbers
- > Recursion and the Fixed-Point Combinator
- > The typed lambda calculus
- > The polymorphic lambda calculus
- > **Other calculi**

Featherweight Java

<p>Syntax:</p> <pre> CL ::= class C extends C {C f̄; K M̄} K ::= C(C f̄) {super(f̄); this.f̄ = f̄;} M ::= C m(C x̄) {return e;} e ::= x e.f e.m(ē) new C(ē) (C)e </pre> <hr/> <p>Subtyping:</p> $C <: C$ $\frac{C <: D \quad D <: E}{C <: E}$ $\frac{CT(C) = \text{class } C \text{ extends } D \{ \dots \}}{C <: D}$ <hr/> <p>Computation:</p> $\frac{fields(C) = \bar{C} \bar{f}}{(new C(\bar{e})) . f_i \rightarrow e_i} \quad (\text{R-FIELD})$ $\frac{mbody(m, C) = (\bar{x}, e_0)}{(new C(\bar{e})) . m(\bar{d}) \rightarrow [\bar{d}/\bar{x}, new C(\bar{e})/this]e_0} \quad (\text{R-INVK})$ $\frac{C <: D}{(D)(new C(\bar{e})) \rightarrow new C(\bar{e})} \quad (\text{R-CAST})$	<p>Expression typing:</p> $\Gamma \vdash x \in \Gamma(x) \quad (\text{T-VAR})$ $\frac{\Gamma \vdash e_0 \in C_0 \quad fields(C_0) = \bar{C} \bar{f}}{\Gamma \vdash e_0 . f_i \in C_i} \quad (\text{T-FIELD})$ $\frac{\Gamma \vdash e_0 \in C_0 \quad mtype(m, C_0) = \bar{D} \rightarrow C \quad \Gamma \vdash \bar{e} \in \bar{C} \quad \bar{C} <: \bar{D}}{\Gamma \vdash e_0 . m(\bar{e}) \in C} \quad (\text{T-INVK})$ $\frac{fields(C) = \bar{D} \bar{f} \quad \Gamma \vdash \bar{e} \in \bar{C} \quad \bar{C} <: \bar{D}}{\Gamma \vdash new C(\bar{e}) \in C} \quad (\text{T-NEW})$ $\frac{\Gamma \vdash e_0 \in D \quad D <: C}{\Gamma \vdash (C)e_0 \in C} \quad (\text{T-UCAST})$ $\frac{\Gamma \vdash e_0 \in D \quad C <: D \quad C \neq D}{\Gamma \vdash (C)e_0 \in C} \quad (\text{T-DCAST})$ $\frac{\Gamma \vdash e_0 \in D \quad C \not<: D \quad D \not<: C \quad \text{stupid warning}}{\Gamma \vdash (C)e_0 \in C} \quad (\text{T-SCAST})$ <p>Method typing:</p> $\frac{\bar{x} : \bar{C}, this : C \vdash e_0 \in E_0 \quad E_0 <: C_0 \quad CT(C) = \text{class } C \text{ extends } D \{ \dots \} \quad \text{override}(m, D, \bar{C} \rightarrow C_0)}{C_0 m (\bar{C} \bar{x}) \{return e_0;\} \text{ OK IN } C}$ <p>Class typing:</p> $\frac{K = C(\bar{D} \bar{g}, \bar{C} \bar{f}) \{super(\bar{g}); this.f̄ = f̄;\} \quad fields(D) = \bar{D} \bar{g} \quad \bar{M} \text{ OK IN } C}{\text{class } C \text{ extends } D \{ \bar{C} \bar{f}; K \bar{M} \} \text{ OK}}$
--	--

Used to prove that generics could be added to Java without breaking the type system.

Igarashi, Pierce and Wadler,
"Featherweight Java: a minimal core calculus for Java and GJ",
 OOPSLA '99
doi.acm.org/10.1145/320384.320395

Featherweight Java is an object calculus inspired by the lambda calculus. Most of the rules have a precondition (over the horizontal line) and a conclusion (below the line). There are rules for dynamic semantics (reductions) and static semantics (type rules).

The β reduction rule is called R-INVK. It says, if method m in class C takes an argument array x and has a body e_0 , then an invocation of a method m with argument array d on an object $(\text{new } C(e))$ reduces to e_0 , with arguments d replacing x and $(\text{new } C(e))$ replacing `this`. (Pretty much like in the lambda calculus.

Featherweight Java is especially interesting because (1) it was used to prove some defects in the Java type system, and (2) it showed that the type system could be extended to incorporate generics without breaking existing code!

Other Calculi

Many calculi have been developed to study the semantics of programming languages.

Object calculi: model *inheritance and subtyping* ..

— lambda calculi with records

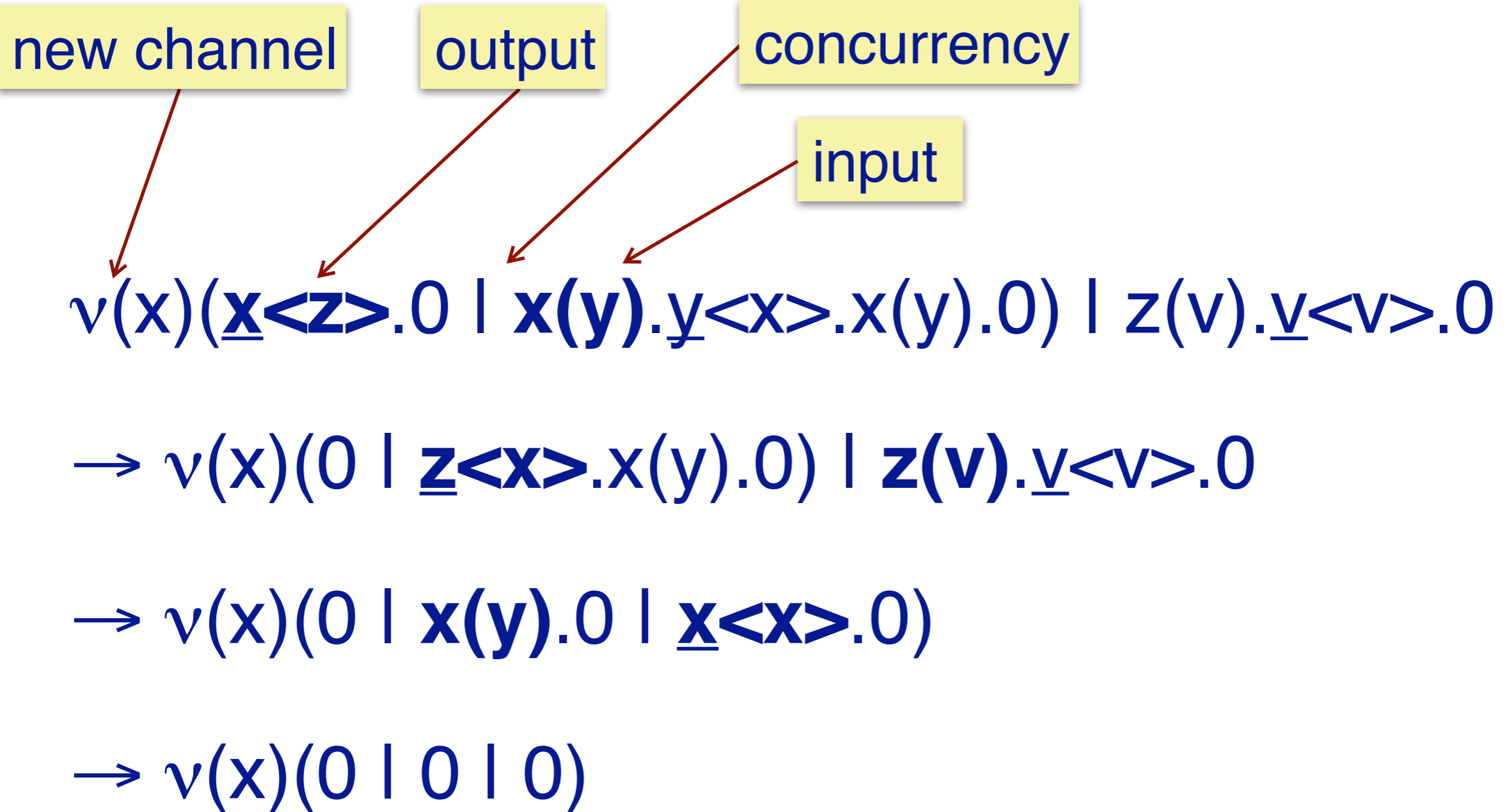
Process calculi: model *concurrency and communication*

— CSP, CCS, pi calculus, CHAM, blue calculus

Distributed calculi: model *location and failure*





— ambients, join calculus

A quick look at the π calculus








1. The value z is output along channel x , to the process $x(y) \cdot \underline{y} \langle x \rangle \dots$, causing y to be bound to z .
2. The value x is output along channel z to the receiving process at the right, $z(v) \cdot \underline{v} \langle v \rangle \cdot 0$, but since x is bound by the scope of $\nu(x) (\dots)$, this forces the receiving process to migrate within the scope of $\nu(x) (\dots)$. Since there is no free x in the migrating process, no variable needs to be renamed.
3. Finally, x is output along channel x , binding it to y . The subsequent processes are all dead (0).

What you should know!

-  *Why isn't it possible to express recursion directly in the lambda calculus?*
-  *What is a fixed point? Why is it important?*
-  *How does the typed lambda calculus keep track of the types of terms?*
-  *How does a polymorphic function differ from an ordinary one?*

Can you answer these questions?

-  How would you model negative integers in the lambda calculus? Fractions?*
-  Is it possible to model real numbers? Why, or why not?*
-  Are there more fixed-point operators other than Y ?*
-  How can you be sure that unfolding a recursive expression will terminate?*
-  Would a process calculus be Church-Rosser?*



Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:



Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.



ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

<http://creativecommons.org/licenses/by-sa/4.0/>