

Programming Language.

Exams come all from Lecture.

What is PL?

- A language for instructing a computer to perform computations.
- also used to express algorithms to other people.

What are [necessary] properties of a PL?

- unambiguous
- implementable sufficiently powerful to
- Turing Complete. (Express any computable functions)

What are [desirable] properties of PLs?

- High level (?)
- system independent (Not tied to any OS or hardware) ^{particular}
- concise.
- Preadable.
- Strong Mathematical foundation (program verification)
- Reusability
- Efficiency (?)
- Exception handling
- Rich set of operators

"High level" vs "Low level" PLs.

Low level PL - constructs of the language reflect the computer.

(Assembly Language Programming) [hardware]

- writing to memory locations
- testing bits and jumping
- perform arithmetic instructions

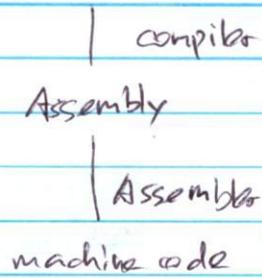
assembly code

move r1, \$x
add r1, r2
jump too

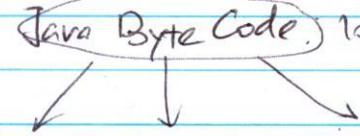
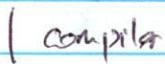
machine code (binary code)

00101101
10010001

C/C++



Java



still consider
low level

High level PL - not low level

- reflects a computational model other than the computer hardware
- mathematical functions
- logic languages - 1st order logic.
- set theory
- :

About mathematical functions:

C = Imperative language

- variables are just memory locations
- assignment statements. write data to memory locations
- loop + jumps for. directing execution over assignment statement

```
int p=1  
for (int i=1, i<=n, i++)  
    p = p * i;
```

} factorial.

ML = functional language, declared.

→ there's no loop

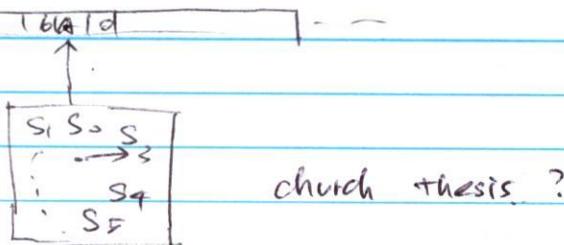
- variables are not memory locations
- no assignment statements that overwrite variables
- no loops, only recursion

$n! = \begin{cases} 1 & \text{if } n = 0 \\ n \times (n-1)! & \text{otherwise.} \end{cases}$

$\text{fun fac } 0. = 1$

| $\text{fac } n = n \times \text{fac}(n-1)$

Turing Machine.



A language is **Turing Complete** if it is equivalent in power to a **universal turing machine**.

↳ UTM is a Turing Machine that can emulate any other TM.

How to define a PL.

Syntax - rules governing the organization of symbols (letters, space, digits, etc) in a valid program.
- defined by grammars

Semantics - give meaning to valid sequence of symbols
static / dynamic

Grammars for defining syntax

Regular Expression:

- used for defining the words (names, keywords, numeric literals) that can be used

CFG - context free grammar

- used to define how words are composed to form expressions, expressions are composed to form statements and so on,

1.2b.

Recitation.

phase of compiler.

- lexer (Token, scanner ...)

- Parser. 1. Take all lexer together and build a structure base on it!

Regular Expression.

$a b^*$.

P_1 is a RE

P_2 is a RE.

concatenation $P_1 P_2$ ab

alternation $P_1 | P_2$ a|b

Kleene star. P_1^* . e,a,aa,aaa

这里 a,b 和 题目不一样

language 'A set of. valid strings'

$a(b|c)^*$

$a b^*$

$a(b|c)^*$

integer - digit *

number - integer | real

real - digit * . digit * digit * In case of different Language
[A-Z] [a-zA-Z]

String only contain 'a' and 'B'

~~b~~ $b|ab|b^*|bb|b^*?$

Only ~~b~~ > b : abba * ba * .

drawback: nesting, palindrome.

CFG: Set of. Non-terminal symbols (do not appear in the final)

set of. terminals (do appear in the final)

start. symbol

Productions.

PF $\xrightarrow{*}$ CFG
CFG $\xrightarrow{*}$ PF

eg: $A \rightarrow 0A1|1$ e. recursive definition.

$A \rightarrow 0A1$ - 00A11

Use CFG to represent: strings contain only 0's and 1's

$A \rightarrow \epsilon | 0A1A | 1A0A$.

Lecture. Syntax (continued)

semantics

- static semantics = rules

governing the use of. types and declarations of names

- Dynamic semantic: defines the execution behavior of.
each valid. construct. (what is +/loop mean?)

Compiler - translates from one PL to another.

↑ Language translate from : source (java, C) ^{human written}
doesn't execute. --- to : target. (Assembly, Machinecode, byte) ^{code}

Interpreter - executes the program

output is the result of running the program.

Phases of compilation

lexical analysis ("lexer")

- forms words (tokens) from sequences of characters.

for (\Rightarrow "for" (keyword)

xyz \Rightarrow "xyz" (Identifier)

1.629 + \Rightarrow 1.629 (Num)

Parsing (aka "syntactic analysis". ("parser"))

- forms sentences from sequences of tokens

- expressions, statements, declarations, functions, modules, programs

sts

Type checking - ensures correct use of types and correct declarations
and use of names

DS

Code generation - (done by code generator)

- produces the target. code.

DS

Optimization - improves the generated code.

Grammars used to define the syntax of PL's

① Regular Expressions

- used to define the valid words in the PL

P.E

Notation



a.

Set

{a}

P₁|P₂ - where P₁|P₂ are P.E's

P₁|P₂ \rightarrow P₁ ∪ P₂

P^{*}, (0 or more concatenations) \rightarrow {the set resulting from zero or more concatenations of strings from P}

S₁, S₂ | S₁ is a string from P₁

S₂ is a st ... P₂}

Alphabet $\{a, b\}$

Regular Operator: $|$, $*$, $()$, concatenation

(cg)

a	$\{a\}$
ab	$\{ab\}$
calbib	$\{ab, bb\}$
$(ab)^*$	$\{\epsilon, a, b, ab, aa, bb, ba, aab, \dots\}$

for | while | let | do

(0|1|2|3|4|5|6|7|8|9)

Shortcut [0-9]

[0-9][0-9]*

wrong $[0-9]^*, [0-9]^* \vdash \epsilon \cdot \epsilon$

$[0-9][0-9]^*, [0-9]^*[0-9] \quad$ you can have 0.0

② Context Free Grammar

consists of:

Non-Terminal symbols

- one of which is designated the start non-terminal

Terminal symbols

Set of rules of the form

$N \rightarrow S \leftarrow$
non-Terminal symbol string of symbols (Term or NonTerm) and

$S \rightarrow aSa$ Non-Terminal: S, B

$S \rightarrow B$ Terminal: a, b

$B \rightarrow bBb$

$B \rightarrow \epsilon$ \in empty string

A CFG defines the set of all strings of ~~non~~-terminals that can be derived from the start non-terminal

- repeatedly applying the rules in the CFG, replacing a non-terminal in a string by the right hand side of a corresponding rule.

derivation:

$$S \Rightarrow aSa \Rightarrow aaSa \Rightarrow aaBa \Rightarrow aabBbaa \Rightarrow aabbbaa$$

$a^n b^m a^n$

CFG defines arithmetic.

$$\text{Exp} \rightarrow \text{Exp} + \text{Exp}$$

$$\text{Exp} \rightarrow \text{Exp} * \text{Exp}$$

$$\text{Exp} \rightarrow \text{ID} \quad (\text{recall possible identifier and numbers})$$

$$\text{Exp} \rightarrow \text{NUM}$$

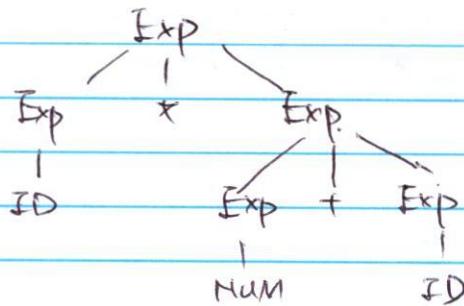
- also written as

$$\text{Exp} \rightarrow \text{Exp} + \text{Exp} \mid \text{Exp} * \text{Exp} \mid \text{ID} \mid \text{NUM}$$

Derivation:

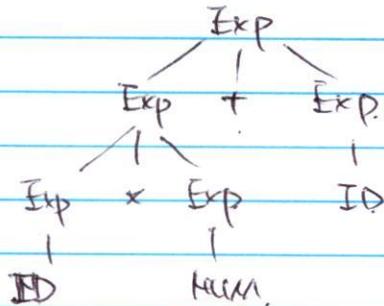
$$\begin{aligned} \text{Exp} &\Rightarrow \text{Exp} * \text{Exp} \Rightarrow \text{ID} * \text{Exp} \Rightarrow \text{ID} * \text{Exp} + \text{Exp} \Rightarrow \text{ID} * \text{NUM} + \text{Exp} \\ &\Rightarrow \text{ID} * \text{NUM} + \text{ID} \Rightarrow x * 3 + y \end{aligned}$$

Parse Tree - graphical representation of a derivation.



Ambiguous Grammar

- allows two parser trees for the same string



How to fix

① Rewrite grammar

② Setting up ~~rules~~ rules

$$\text{Exp} \rightarrow \text{Term} + \text{Term} \mid \text{Term}$$

$$\text{Term} \rightarrow \text{Factor} * \text{Factor} \mid \text{Factor}$$

$$\text{Factor} \rightarrow \text{ID} \mid \text{NUM} \mid (\text{Exp})$$

$a^n b^m c^n$ No way to write in CFG

Scoping of a name portion of the program in which the name is visible.

Block - syntactic construct that defines a scope.

```
{ int x  
}
```

```
void f( inty ) {  
    }
```

Block Structure Language -

a PL in which blocks can be nested

- including nesting functions

In a block structure language, a variable reference is resolved to the corresponding declaration in the innermost surrounding block.

static
scoping

```
x: integer;  
procedure f();
```

begin

```
x := x+1;
```

end

```
procedure g();
```

begin

```
x: integer = 7;
```

begin

```
f();
```

end.

Q: What x? Depends on language is in what kind of scoping.

Dynamic Scoping

Predation:

Static/Dynamic Semantic Definition 考试会考背下来

e.g. end in 'a' \Rightarrow set of {a, b}

$S \rightarrow aS \mid bS \mid a$

e.g. even a's \Rightarrow set of {a, b}

$S \rightarrow aSa \mid \epsilon \mid bS$

e.g. 5. Give CFG read x

read y

prod = x^*y

write prod

program \rightarrow start-list.

$x^*(\rightarrow)^* + q$

start-list \rightarrow state start-list

↑

state \rightarrow read id | write Expr | id = Expr

Ep \rightarrow id

2/8.

Lecture.

static scoping: The body of a function is evaluated in the environment in which the function is defined

- environment: mapping of names to values or memory locations.

Dynamic scoping:

called

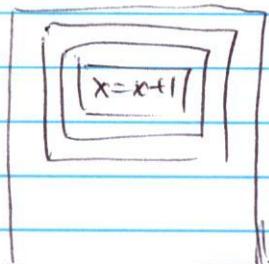
* local declaration always precedes *

static scoping

compiler figures it out

dynamic scoping.

call chain



Compile time.

f g h i k

← looking backwards

run time

dynamic scop
g>f>I

```

x: integer = 7;
procedure f( procedure h() );
begin
  h();
end
procedure g();

```

Q
If call g();
starts.

static scoping

```

x: integer := 20;
procedure I();
begin
  x := x + 1;
end;
begin
end f(I)

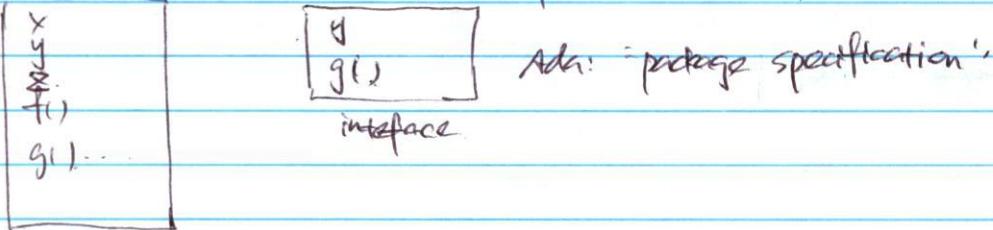
```

Ada

- modularity ("packages")
- concurrency

Module - a collection of code

- has a restricted interface to the rest of the program
to code outside of the module



module body } Ada: "package body"

Concurrency - the order in which events may happen,
is unknown

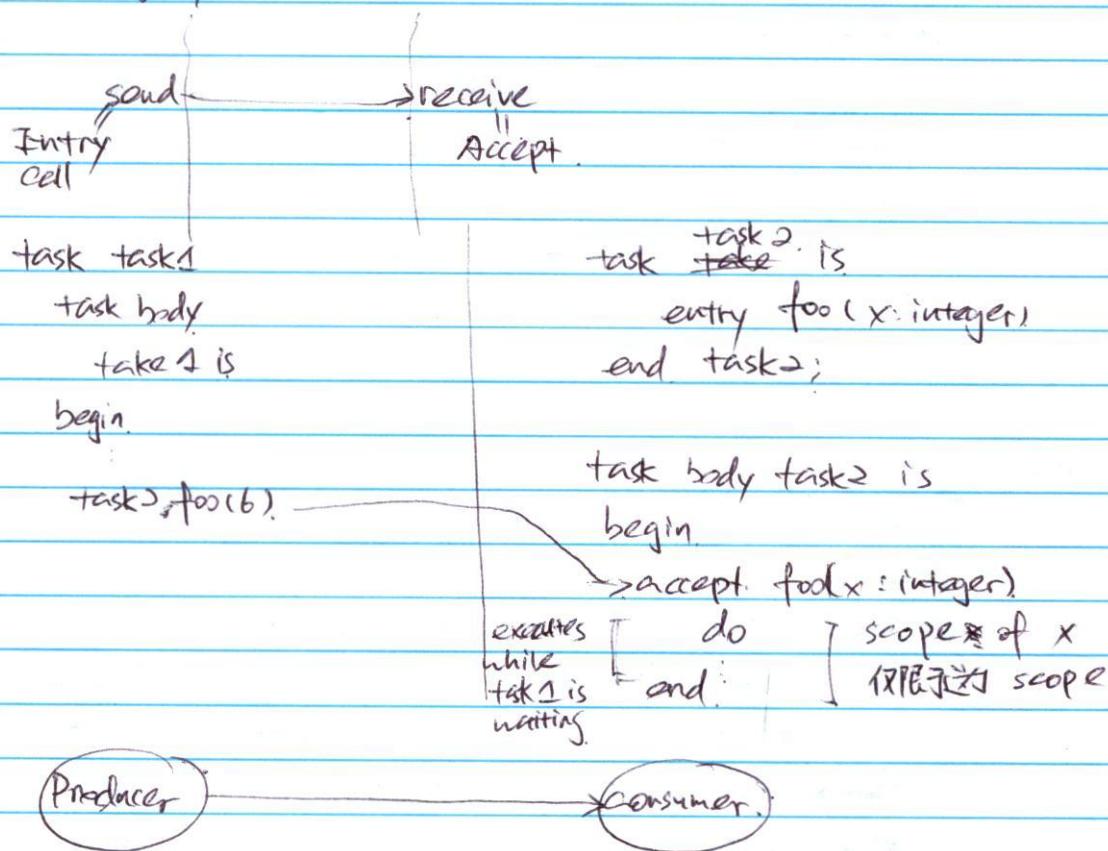
- no assumption can be made about the
order of events

Ada uses thread-based concurrency

- a thread is a sequence of operations (instructions).
- two threads are concurrent if no assumptions can be made about the order of operations in one thread relative to the other thread.

Ada: A thread is called a Task

like Msg synchronization in OS.



Task Type

task type MyTask is
entry Go;
end.

task body MyTask is
begin
:
end

T1: MyTask
T2: MyTask
A: array(1..100) of MyTask
T1: Go
T2: Go
for i in 1..100 loop
A(i) := Go;
end loop;

Parameter Passing

Math $f(x) = x^2 + 2$
let $y = 10$.
 $f(y) = ? = 100 + 2 = 102$

} don't need to think about
how y is "passed" to x

In imperative languages, variables are memory locations.

For parameter passing mechanisms:

① Pass by value

- the value of the actual parameter is copied to the formal parameter.

void f (int x)

{

$x = x + 1;$

}

int y = 3
f(y) ← actual parameter.
print(y)

x [4]

y [3]

when function
returns, x gets away. → copy

② Pass by reference

- the address of the actual parameter is passed into the formal parameter.
- any reference to the formal parameter follows the address back to the actual parameter

void f (int x)

{

$x = x + 1;$

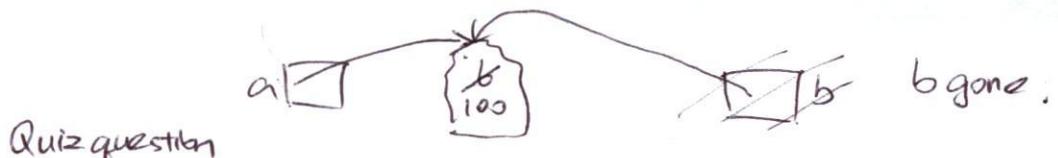
}

int y = 3

f(y)
print(y)

x [4]

y [3] 4



Quiz question.

Class A {int x;}

A a = new A();

a.x = 6;

g(a);

print(a.x); // 100. ~~Java is pass by value~~

void g(A b),

{ b.x = 100;

}

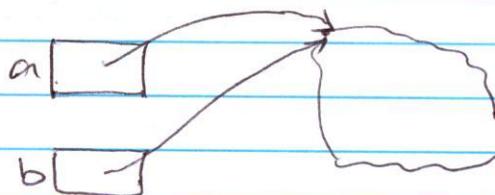
Object has pointer semantics.

- - - connecting to Array. - - -

Java's objects have "pointer semantics"!

- the "value" of an object is its address.

A a = new A();



A b = a

~~This is a function now~~
~~1 & 2 is pass by value~~ ~~3 is pass by reference~~ ~~If b point to a~~ ~~newA()~~

If Java is

③ Pass by value result.

= copy in., copy out!

- Upon procedure call, copy the values of the actual params to the corresponding formal parameters
- upon procedure return copy the values of the formal parameters back to the actual params

结果和参数是否一样了？

int y	pass by reference	pass by result.
void f (int x)	dependent.	independent
{		
y = y + 1	y []	y [] 4
	x []	x [] 4
x = x + 1		* back to x.
}		
	print 5.	print 4.
y = 3		
f(y)		
print(y);		

⑦ Pass by name (Not that efficient)

- The semantics of a function call is defined by textual substitution.
- "Alg copy Rule"

Step 1: Replace a function call with the body of the function definition,

renaming variables at the call site to avoid name conflicts.

Step 2: Replace all occurrences of the formal parameters in the function body with the actual parameters. - textual replacement

- rename local variables in the function body to avoid name conflicts.

procedure f(x: integer, y: integer);

begin

y = x + 1;

y = y + 1;

end.

z: integer := 15;

w: integer;

begin

step 1

step 2

$z := z + 1; \quad \left\{ \begin{array}{l} y = x + 2 \\ f(z + 3, w) \end{array} \right. \Rightarrow w = (z + 3) * 2$

$y = y + 1 \Rightarrow w = w + 1$

print(w);

$\Rightarrow 39$

end;

Prediction: ① (in Integer) recommended to add

Ans: By default, (Integer) is taken as (in Integ).

② A := Procedure ; X differ from other languages:

A := Function ; Don't need ()

③ Not that Case Sensitive.

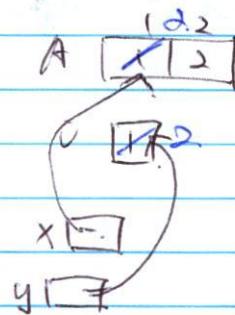
④ - ads. - specification

- adb - body

7/22.

Q. what's the difference between pass by reference and pass by name?

Q2. pass by reference 2



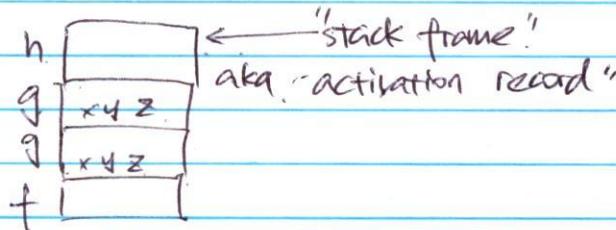
Pass by name. 1

$$g(A(i), i) \rightarrow \cdot \vec{v} = \vec{v} + 1; \quad 4$$

A [1 2 3 4] $A(i)$
 i [1 2]

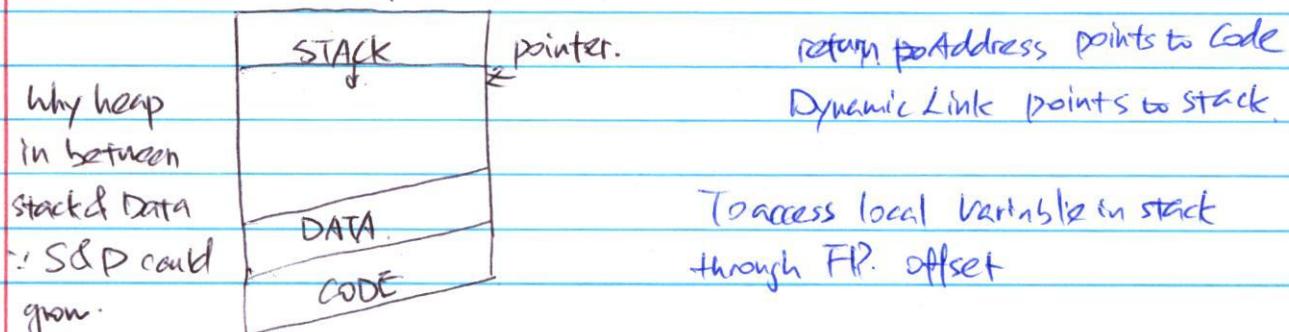
$$A_{ij} = A_{i,j} * 2.$$

Allocating function calls on a stack
 $f \sim g \sim h \sim h$



Memory Organization on the x86

Address Space for a program



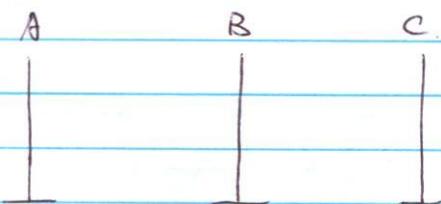
Functional Languages

- Variables denote value, not memory locations
- Recursion must be used instead of loops for repetition
- Functions are values
 - "first class objects"
- functions that operate on other functions are called
"higher order functions"

Recursive Thinking

- ① Solve the base case.
 - i.e. for the smallest possible problem size
 - ② Assume that the function you are writing works correctly on size $N-1$.
 - ③ Under assumption ②; construct the solution for size N .
- Example: Factorial
- ① Base case $\text{Fac}(0) = 1$
 - ② Assumption $\text{Fac}(n-1) = (n-1)!$
 - ③ $\text{Fac}(n) = \text{Fac}(n-1) * n$.

Towers of Hanoi



Move N disks from A to B

- ① Base case. $N=1$
 - Move the disk from A to B
- ② Assumption:

Can move $N-1$ disks from one peg to another, using the third peg as temporary

move($N-1$, FROM, TO, TEMP)

- ③ Define move (N , FROM, TO, TEMP)
as.

move (N-1, FROM, TEMP, TO)

move 1 disk from FROM to TO

move (N-1, TEMP, TO, FROM)

March 27 - Midterm. (1 hour). How you learned in the class.

- LISP is dynamic scoping but ^{Scheme} ~~still~~ fix it.

Schema - a dialect. of Lisp.

- statically scoped
- dynamically typed

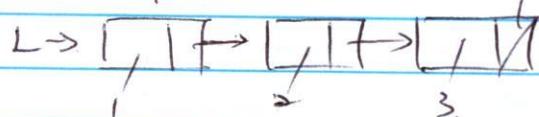
why call car & cdr?

'(1 2 3)

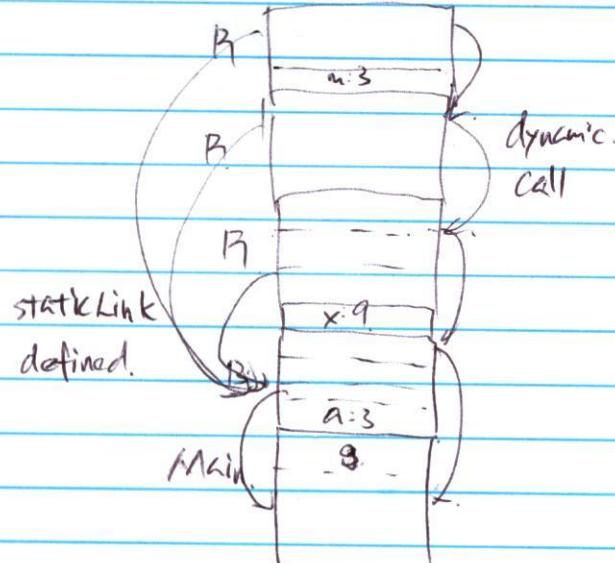
context of Decoder Register

context of Address Register.

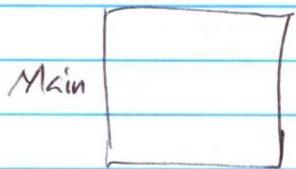
NIL



(null? L) → testing for the empty list



Example 2



e.g. length of list

- (1) base case. empty list. 0.
- (2) Assumption. n-1 length has been found.
- (3) ($+ 1$ Len(cdr L))

(define len L

```
(cond ((null? L) 0)
      (else (+ 1 (len (cdr L))))))
```

3/8.

TM Completeness

Syntax Semantics

Scoping

Adh tasks package

The λ -calculus

- syntactic model of computation
- based on transformations applied to expressions

3/22

Syntax of the λ -calculus

$$E \rightarrow C \mid x \mid E_1 E_2 \mid \lambda x. E$$

Terminology

"Free" occurrence of a variable

- think of it as a "non-local" variable

$$\begin{array}{ccc} x \mid x \mid y & \lambda x. \mid x \mid & \\ \uparrow \text{free} & & \uparrow \text{not free = bound} \\ (\lambda x. (\lambda x. x)) & & \end{array}$$

Rules for "free"-ness

x is free in x , but not in any other variable y .

or any constant b

* x is free in $E_1 E_2$ iff x is free in E_1 or in E_2

* x is free in $(\lambda y. E)$ iff x is free in E and y is not the same as x

$$\lambda y. x$$

 $\uparrow \text{free}$

Substitution

Notation: $E[M/x]$

- E with all free occurrences of x replaced by M .

- example: $(x \cdot y) [(\lambda z. z)/y]$

$$\Rightarrow (x \cdot (\lambda z. z))$$

Rules for substitution

$$x[M/x] = M$$

$$c[M/x] = c$$

$y[M/x] = y$, where y is different from x .

$$(E_1 E_2)[M/x] = (E_1[M/x] E_2[M/x])$$

* $(\lambda x. E)[M/x] = \lambda x. E$ (by definition)

* $(\lambda y. E)[M/x] =$

{ $(\lambda y. E[M/x])$ if y does not appear free in M (no conflict)

↓ $(\lambda z. E[z/y])[M/x]$ otherwise, where Z is a new variable.

now → $(\lambda y. +x 1)[y/x]$

$$x \Rightarrow (\lambda z. +x 1)[y/x] \Rightarrow (\lambda z. +y 1).$$

Conversion Rules

L conversion 不帶 y 的時候.

If y is not free in E , then $(\lambda x. (x)) \Leftrightarrow (\lambda y. (y))$

$$(\lambda x. E) \Leftrightarrow (\lambda y. E[y/x]) \Leftrightarrow (\lambda y. (y))$$

* B conversion (pass by name)

$$(\lambda x. E) M \Leftrightarrow E[M/x]$$

eg. $(\lambda x. +x x)(+23)$

$$\xrightarrow{B} +(+23)(+23)$$

η conversion

$$(\lambda x. E x) \Leftrightarrow E$$

If x is not free in E .

church encoding

$$\text{False } \lambda a. \lambda b. b$$

$$\text{True } \lambda a. \lambda b. a.$$

3/29.

$$\begin{aligned} & \text{not free.} \\ & (+ x ((\lambda x . x) y) z) \quad [(\lambda x . x) / x] \\ \Rightarrow & (+ (z z) ((\lambda x . x) y) z) \end{aligned}$$

β -reduction is β conversion in \Rightarrow direction
 $(\lambda x . E) M \Rightarrow E[M/x]$

- Numbers + operators can be encoded as expressions in the λ -calculus and use β -reduction to evaluate them.

But, we will take a short cut: δ -reduction - specifies the behavior of all operators.

$$+ 4 5 \xrightarrow{\delta} 9$$

$$\text{if true } E_1 E_2 \xrightarrow{\delta} E_1$$

$$\text{if false } E_1 E_2 \xrightarrow{\delta} E_2$$

Any expression that β - or δ reduction can be applied to is called a reducible expression = redpx,

Any expression that cannot be reduced is in "Normal Form".
b. $(\lambda x . x)$

Q: Does every expression have a normal form?

No. $(\lambda x . x)(\lambda x . x) \xrightarrow{\delta} (\lambda x . x)(\lambda x . x) \Rightarrow \dots$

Recursion using the Y combinator

- has a special property: the fixpoint of f .

$$Y f \Leftrightarrow f(Y f)$$

- for this reason, Y is also called a "fixed point" combinator.

The fixpoint of a function f is the value z such that $f(z) = z$

e.g. the fixpoint of $f(x) = 2 * x$ is 0
 $\vdash f(0) = 0$

use recursion $yf = f(yf)$

Factorial

$$y(\lambda f. \lambda x. \text{if}(=x 0) 1 (*x(f(-x 1)))$$

FAC

$$\text{FAC } 3: (y(\lambda f. \lambda x. \dots)) 3$$

$$(\underline{\lambda f. \lambda x. \dots})(\underline{y(\lambda f. \lambda x. \dots)}) 3 \quad yf = f(yf)$$

$$\xrightarrow{\beta} \lambda x. \text{if}(=x 0) 1 (*x(x(y(\lambda f. \lambda x. \dots)(-x 1)))) 3$$

$$\xrightarrow{\beta} \text{if}(=3 0) 1 (*3(y(\lambda f. \lambda x. \dots)(-3 1)))$$

$$\xrightarrow{\beta} \dots \Rightarrow (*3(\underbrace{y(\lambda f. \lambda x. \dots)}_{\text{FAC}}) 2 1)$$

y can be expressed in the λ -calculus as:

$$(\lambda h. (\lambda x. h(xx))(\lambda x. h(xx)))$$

$$yf = (\lambda h. (\lambda x. h(xx))(\lambda x. h(xx))) f.$$

$$\xrightarrow{\beta} (\lambda x. f(xx))(\lambda x. f(xx))$$

$$\xrightarrow{\beta} f. (\lambda x. f(xx))(\lambda x. f(xx))$$

$$\xrightarrow{\beta} f (\lambda h. (\lambda x. h(xx))(\lambda x. h(xx)) + f)$$

$$= f(yf)$$

Conversion
in the ϵ direction

Orders of Evaluation: The order in which redexes are chosen to reduce. δ -reduction

$$(\lambda x. +xx)(+\overline{3}4)$$

β -reduction.

Normal Order Reduction

Always choose the leftmost outermost redex to reduce

- call the function before evaluating the arguments

Applicative Order Reduction

Always choose the leftmost innermost redex to reduce

- Evaluate arguments before calling the function.

$$(\lambda y.z)(\lambda x.xx)(\lambda x.xx)$$

Normal: 3.

terminates

$$A. : (\lambda y.z)(\lambda x.xx)(\lambda x.xx)$$

Not terminates

Church-Rosser Theorem 1:

- All terminating reduction sequences for an expression will result in the same normal form.

Church-Rosser Theorem 2:

- If any reduction sequence terminates, then normal order reduction will terminate

4/12. $x+yz$ (op+) (x,y) infix op only takes 2 para tuples.

ML cont. fun f. (op <) $x + y = \frac{x \leftarrow y}{\text{anytype type}}$

$\frac{((d * \beta) * y)}{c} \rightarrow \frac{\alpha}{x} \rightarrow \frac{\beta}{y} \rightarrow \frac{f}{\text{result}}$ don't have to be the same type.

(Define its own infix operator)

$3 *=*= 4$

(op *=*=) (3, 4)

infix ++ :

texx infix ff. 不是 fun (op++) a b = a+b;

v. fun a ++ b = a+b ;

What features make a language "object oriented"?

- ① Encapsulation of data and code into a single structure.
 - fields and methods w/in an object
- ② Inheritance.
 - defines a new type based on an existing type, where the new type can reuse code defined in the existing type.
 - the child type inherits the methods of the parent type, as well as the fields

class vehicle {

 int speed;

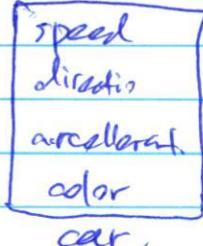
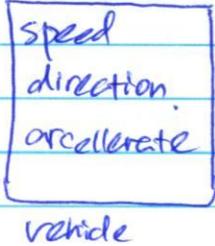
 float direction;

 void accelerate (int delta) { -- }

}

class car extends vehicle {

 int color;



- ③ Subtyping with dynamic dispatch.

"dynamic overloading"

- subtyping: treating a type as if it were another type

- anywhere a type T can be used, a subtype of T can be used.

T_1

T_2

} "T₂ is a subtype of T₁"

$T_1 x = \text{new } T_2()$;

void. $f(T_1, x)$ {
 \dots
 g }

$T_2 \models T_1$

$f(g)$;

- think of subtyping as an "is a" relation.

- an object of type T_2 "is a" object of type T_1 .

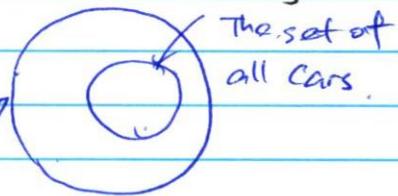
Dynamic dispatch: determining at run time which method to call, based on the actual type of an object, not the declared type.

A type is a set of elements

- therefore operators defined on these elements

class vehicle {
 \dots

The set of
all vehicles



class car extends Vehicle {
 \dots

In the subset interpretation of "subtyping" a subtype of a type T denotes a subset of the set denoted by T .

what if there was subtyping between function types.

- there is, in Scala

(Using ML syntax, but ML does not have this!)

A $\text{fun } f(g: A \rightarrow \text{int}) = g(\text{new } A(1))$

|

if not exist in A.

A is B

B is A.

B $\text{fun. } h(x: B) = x.\text{foo} + 1$

$B \rightarrow \text{int}$

(Q1) $f(h)$ not ok!

You cannot use a $B \rightarrow \text{int}$ where an $A \rightarrow \text{int}$ is expected.

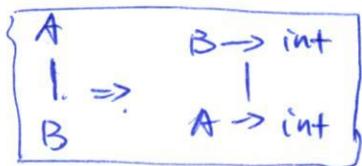
$\text{fun. } f(g: B \rightarrow \text{int}) = g(\text{new } B(1))$

$\text{fun. } h(x: A) = x.\text{foo} + 1$

$A \rightarrow \text{int}$

(Q2) $f(h)$ is ok!

You can use an $A \rightarrow \text{int}$ wherever a $B \rightarrow \text{int}$ is expected.



Function Subtyping is contravariant on the (input) types

$\vdash C$ $\text{fun } f. (g: \text{int} \rightarrow C) = \text{let } y = C = g(6);$
 $|$ $\text{fun } h. (x: \text{int.}) = \text{new } D();$
 $\text{Car. } D$ $\text{int} \rightarrow D.$

$f(h)$ is ok. D is C .

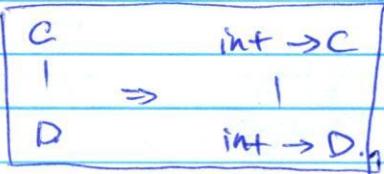
$\text{fun. } f (g: \text{int} \rightarrow D) = \text{let } y = D = g(6);$

$\text{fun } h (x: \text{int.}) = \text{new } C();$
 $|$
 $f(h) \quad \text{int} \rightarrow C$

$f(h)$ is not ok. C is not D

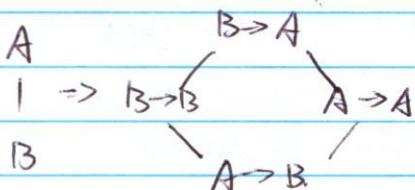
Anywhere a. $\text{int} \rightarrow C$ can be used.

$\text{int} \rightarrow D$ can be used.



Function Subtyping is covariant on the output types

$A \quad C \quad B \rightarrow Q.$
 $| \quad | \Rightarrow \quad |$
 $B \quad D \quad A \rightarrow D.$



7/19

Function Subtyping :

- contravariant on the input type.

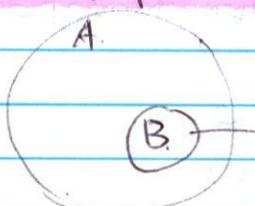
If B is a subtype of A , then $A \rightarrow \text{Int}$ is a subtype of $B \rightarrow \text{Int}$.

- covariant on the output type

If B is a subtype of $\&A$, then $\text{int} \rightarrow B$ is a subtype of $\text{int} \rightarrow A$.

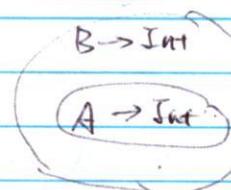
Subset Interpretation of subtyping.

A
|
B



though B might have more in itn.

A $B \rightarrow \text{Int}$
| |
B $A \rightarrow \text{Int}$

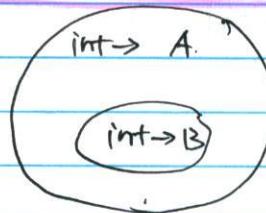


• $A \rightarrow \text{Int}$ denotes the set of all functions that can be applied to an A object and return an int

• $B \rightarrow \text{Int}$ denotes the set of all functions that can be applied to a B object and return an int

Since any function expecting an A can be applied to B , any function in $A \rightarrow \text{Int}$ is also in $B \rightarrow \text{Int}$. So $A \rightarrow \text{Int}$ is a subset of $B \rightarrow \text{Int}$

A $\text{int} \rightarrow A$
| |
B $\text{int} \rightarrow B$



• $\text{int} \rightarrow B$ denotes the set of all functions that take an int and return a B .

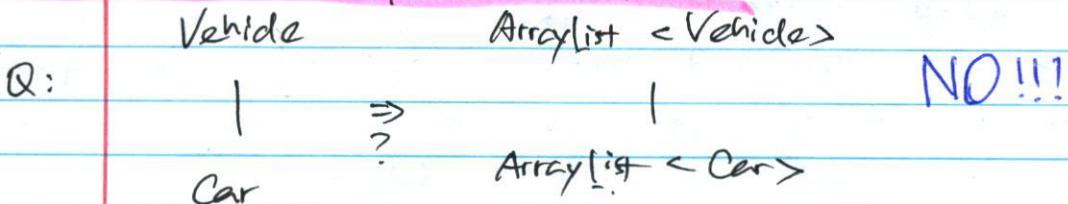
• $\text{int} \rightarrow A$ denotes the set of - - - - -

• Since a B object is an A object, any function returning B also returns A . So any function in $\text{int} \rightarrow B$ is also in $\text{int} \rightarrow A$, so $\text{int} \rightarrow B$ is a subset of $\text{int} \rightarrow A$.

Parameter.
type.

interface $I < T \rangle$ Entry $< T, T \rangle$ twice ($*T$ value).

Java Generic does not implement Subtyping.



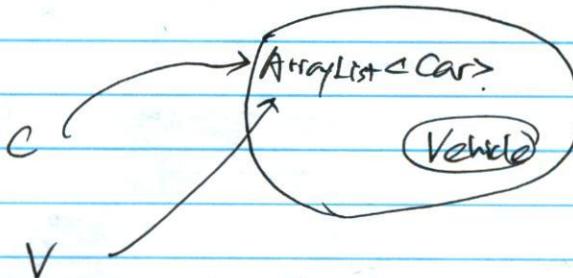
Assumption: subtype is allowed.

$\text{ArrayList } < \text{Car} \rangle c = \text{new ArrayList } < \text{Car} \rangle$

$\text{ArrayList } < \text{Vehicle} \rangle V = c$

$V.add(\text{new Vehicle})$;

$\text{Car } x = c.get(0)$;



void AddVehicle ($\text{ArrayList } < \text{Vehicle} \rangle V$) {
 $V.add(\text{new Vehicle})$;

$\text{ArrayList } < \text{Car} \rangle c = \text{new ArrayList } < \text{Car} \rangle()$

$\text{AddVehicle}(c);$

⇒ So $\text{ArrayList } < \text{Car} \rangle$ is not a subtype of $\text{ArrayList } < \text{Vehicle} \rangle$

Q: Is $\text{ArrayList } < \text{Vehicle} \rangle$ a subtype of $\text{ArrayList } < \text{Car} \rangle$? NO!!!

$\text{ArrayList } < \text{Vehicle} \rangle V = \text{new ArrayList } < \text{Vehicle} \rangle$

$\text{ArrayList } < \text{Car} \rangle c = V$;

$\text{Car } x = c.get(0)$;

* No subtyping between instances of A generic *

Defining a function that operates on different types of ArrayList.

```
void f (ArrayList<Object> A) {
```

```
    for (Object e = A) ?
```

```
        S.O.P (e)
```

```
X A.add (new Object ())
```

```
}
```

↑ 这里不行。应该 pass in Car. Object 和？不匹配。

和前面一样：错误。

Solution: #1?

syntax matches any kind
of list.

Vehicle

```
void f (ArrayList<? extends Vehicle> V)
```

```
|
```

```
{ for (Vehicle e = V)
```

Can't stick anything into List.

Car

```
    print (e.speed);
```

V ↪ C

```
|
```

```
}
```

poche

```
void g (ArrayList<? super Car> C)
```

```
{
```

```
C.add (new Car ());
```

super ↑
upper }

extends ↓
lower .

In bounded.java. class C < T extends A > {

T x

ArrayList<T> f ()

}

class C < T extends A >

extends D < T > D is parent..

implements I < T > { --- }

Scala Notes

? $a+b$ $\rightarrow a.f(b)$
? $a.+b)$ $\rightarrow a+b$

def f(c: Collection[A]) {
 c.add(new A())

}

var cb: Collection[B]
f(cb) Wrong : C[A]
 | x.
 C[B]

def f(c: Collection[B]) {
 c.add(new B())

}

var ca: Collection[A]
f(ca) then works

A	C[B]
	contra case
B	C[A]

def g(c: Collection[A]) {
 val a: A = c.get(0)

}

var cb: Collection[B] = ---
g(cb)

C[A]	
	covar case
C[B]	

Back to Polymorphism

ML: parametric Polymorphism

DOP: bounded Polymorphism

- subtype
- interface