Programming Languages and Compilers (CS 421)

Elsa L Gunter

2112 SC, UIUC

http://courses.engr.illinois.edu/cs421

Based in part on slides by Mattox Beckman, as updated by Vikram Adve and Gul Agha

Nested Recursive Types

```
# type 'a labeled_tree =
TreeNode of ('a * 'a labeled_tree
list);;
type 'a labeled_tree = TreeNode of ('a
* 'a labeled_tree list)
```

4

Nested Recursive Type Values



Nested Recursive Type Values

```
val ltree : int labeled_tree =
  TreeNode
  (5,
    [TreeNode (3, []); TreeNode (2,
    [TreeNode (1, []); TreeNode (7, [])]);
    TreeNode (5, [])])
```

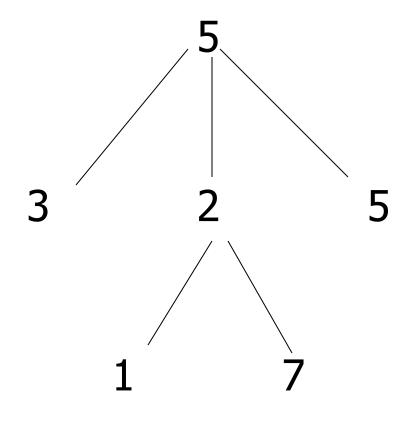


Nested Recursive Type Values

```
Ltree = TreeNode(5)
TreeNode(3) TreeNode(2) TreeNode(5)
          TreeNode(1) TreeNode(7)
```



Nested Recursive Type Values





Mutually Recursive Functions

```
# let rec flatten_tree labtree =
  match labtree with TreeNode (x,treelist)
    -> x::flatten tree list treelist
  and flatten tree list treelist =
  match treelist with [] -> []
   | labtree::labtrees
    -> flatten tree labtree
      @ flatten tree list labtrees;;
```

4

Mutually Recursive Functions

 Nested recursive types lead to mutually recursive functions



- Data types play a key role in:
 - Data abstraction in the design of programs
 - Type checking in the analysis of programs
 - Compile-time code generation in the translation and execution of programs
 - Data layout (how many words; which are data and which are pointers) dictated by type

Terminology

- Type: A type t defines a set of possible data values
 - E.g. short in C is $\{x \mid 2^{15} 1 \ge x \ge -2^{15}\}$
 - A value in this set is said to have type t

 Type system: rules of a language assigning types to expressions



Types as Specifications

- Types describe properties
- Different type systems describe different properties, eg
 - Data is read-write versus read-only
 - Operation has authority to access data
 - Data came from "right" source
 - Operation might or could not raise an exception
- Common type systems focus on types describing same data layout and access methods



Sound Type System

If an expression is assigned type t, and it evaluates to a value v, then v is in the set of values defined by t

- SML, OCAML, Scheme and Ada have sound type systems
- Most implementations of C and C++ do not



Strongly Typed Language

- When no application of an operator to arguments can lead to a run-time type error, language is strongly typed
 - Eg: 1 + 2.3;;
- Depends on definition of "type error"



Strongly Typed Language

- C++ claimed to be "strongly typed", but
 - Union types allow creating a value at one type and using it at another
 - Type coercions may cause unexpected (undesirable) effects
 - No array bounds check (in fact, no runtime checks at all)
- SML, OCAML "strongly typed" but still must do dynamic array bounds checks, runtime type case analysis, and other checks



Static vs Dynamic Types

- Static type: type assigned to an expression at compile time
- Dynamic type: type assigned to a storage location at run time
- Statically typed language: static type assigned to every expression at compile time
- Dynamically typed language: type of an expression determined at run time

Type Checking

- When is op(arg1,...,argn) allowed?
- Type checking assures that operations are applied to the right number of arguments of the right types
 - Right type may mean same type as was specified, or may mean that there is a predefined implicit coercion that will be applied
- Used to resolve overloaded operations



- Type checking may be done statically at compile time or dynamically at run time
- Dynamically typed (aka untyped) languages (eg LISP, Prolog) do only dynamic type checking
- Statically typed languages can do most type checking statically



Dynamic Type Checking

- Performed at run-time before each operation is applied
- Types of variables and operations left unspecified until run-time
 - Same variable may be used at different types



Dynamic Type Checking

- Data object must contain type information
- Errors aren't detected until violating application is executed (maybe years after the code was written)



Static Type Checking

- Performed after parsing, before code generation
- Type of every variable and signature of every operator must be known at compile time



Static Type Checking

- Can eliminate need to store type information in data object if no dynamic type checking is needed
- Catches many programming errors at earliest point
- Can't check types that depend on dynamically computed values
 - Eg: array bounds



Static Type Checking

- Typically places restrictions on languages
 - Garbage collection
 - References instead of pointers
 - All variables initialized when created
 - Variable only used at one type
 - Union types allow for work-arounds, but effectively introduce dynamic type checks



- Type declarations: explicit assignment of types to variables (signatures to functions) in the code of a program
 - Must be checked in a strongly typed language
 - Often not necessary for strong typing or even static typing (depends on the type system)

Type Inference

- Type inference: A program analysis to assign a type to an expression from the program context of the expression
 - Fully static type inference first introduced by Robin Miller in ML
 - Haskle, OCAML, SML all use type inference
 - Records are a problem for type inference

Format of Type Judgments

A type judgement has the form

$$\Gamma$$
 |- exp : τ

- I is a typing environment
 - Supplies the types of variables (and function names when function names are not variables)
 - Γ is a set of the form $\{x:\sigma,\ldots\}$
 - For any x at most one σ such that $(x : \sigma \in \Gamma)$
- exp is a program expression
- $\mathbf{\tau}$ is a type to be assigned to exp
- pronounced "turnstyle", or "entails" (or "satisfies" or, informally, "shows")



Axioms - Constants

 $\Gamma \mid -n : int$ (assuming *n* is an integer constant)

 Γ |- true : bool

 Γ |- false : bool

- These rules are true with any typing environment
- \blacksquare Γ , n are meta-variables



Axioms – Variables (Monomorphic Rule)

Notation: Let $\Gamma(x) = \sigma$ if $x : \sigma \in \Gamma$

Note: if such o exits, its unique

Variable axiom:

$$\Gamma \mid -x : \sigma$$
 if $\Gamma(x) = \sigma$



Simple Rules - Arithmetic

Primitive operators (
$$\oplus \in \{+, -, *, ...\}$$
):
$$\frac{\Gamma \mid - e_1 : \tau_1 \qquad \Gamma \mid - e_2 : \tau_2 \qquad (\oplus) : \tau_1 \rightarrow \tau_2 \rightarrow \tau_3}{\Gamma \mid - e_1 \oplus e_2 : \tau_3}$$
Relations ($^{\sim} \in \{<, >, =, <=, >= \}$):
$$\frac{\Gamma \mid - e_1 : \tau \qquad \Gamma \mid - e_2 : \tau}{\Gamma \mid - e_1 \qquad e_2 : \text{bool}}$$

For the moment, think τ is int

Example: $\{x:int\} | -x + 2 = 3 : bool$

What do we need to show first?

$$\{x:int\} \mid -x + 2 = 3 : bool$$

4

Example: $\{x:int\} | -x + 2 = 3 : bool$

What do we need for the left side?

Example: $\{x:int\} | -x + 2 = 3 : bool$

How to finish?

```
\{x:int\} \mid -x:int \ \{x:int\} \mid -2:int \ \{x:int\} \mid -x+2:int \ \{x:int\} \mid -3:int \ \{x:int\} \mid -x+2=3:bool
```

4

Example: $\{x:int\} | -x + 2 = 3 : bool$

Complete Proof (type derivation)



Simple Rules - Booleans

Connectives

$$\Gamma \mid -e_1 : bool$$
 $\Gamma \mid -e_2 : bool$ $\Gamma \mid -e_1 \&\& e_2 : bool$

$$\Gamma \mid -e_1 : bool$$
 $\Gamma \mid -e_2 : bool$ $\Gamma \mid -e_1 \mid e_2 : bool$

Type Variables in Rules

If_then_else rule:

```
\Gamma \mid -e_1 : \text{bool} \quad \Gamma \mid -e_2 : \tau \quad \Gamma \mid -e_3 : \tau
\Gamma \mid -(\text{if } e_1 \text{ then } e_2 \text{ else } e_3) : \tau
```

- τ is a type variable (meta-variable)
- Can take any type at all
- All instances in a rule application must get same type
- Then branch, else branch and if_then_else must all have same type

4

Function Application

Application rule:

$$\frac{\Gamma \mid -e_1 : \tau_1 \rightarrow \tau_2 \quad \Gamma \mid -e_2 : \tau_1}{\Gamma \mid -(e_1 e_2) : \tau_2}$$

If you have a function expression e_1 of type $\tau_1 \rightarrow \tau_2$ applied to an argument e_2 of type τ_1 , the resulting expression e_1e_2 has type τ_2

Fun Rule

- Rules describe types, but also how the environment \(\Gamma\) may change
- Can only do what rule allows!
- fun rule:

$$\{x : \tau_1\} + \Gamma \mid -e : \tau_2$$

$$\Gamma \mid -\text{fun } x -> e : \tau_1 \to \tau_2$$

Fun Examples

```
\{y : int \} + \Gamma \mid -y + 3 : int \}
 \Gamma \mid -fun y -> y + 3 : int \rightarrow int \}
```

```
\{f: int \rightarrow bool\} + \Gamma \mid -f 2 :: [true] : bool list
 \Gamma \mid -(fun f -> f 2 :: [true])
 : (int \rightarrow bool) \rightarrow bool list
```



(Monomorphic) Let and Let Rec

let rule:

$$\Gamma \mid -e_1 : \tau_1 \quad \{x : \tau_1\} + \Gamma \mid -e_2 : \tau_2$$

$$\Gamma \mid -(\text{let } x = e_1 \text{ in } e_2) : \tau_2$$

let rec rule:

$$\{x: \tau_1\} + \Gamma \mid -e_1:\tau_1 \{x: \tau_1\} + \Gamma \mid -e_2:\tau_2$$

 $\Gamma \mid -(\text{let rec } x = e_1 \text{ in } e_2):\tau_2$

Example

Which rule do we apply?

```
|- (let rec one = 1 :: one in let x = 2 in fun y \rightarrow (x :: y :: one)) : int \rightarrow int list
```

Example

```
(2) {one : int list} |-
Let rec rule:
                             (let x = 2 in
                         fun y -> (x :: y :: one))
{one : int list} |-
(1 :: one) : int list
                             : int \rightarrow int list
 |- (let rec one = 1 :: one in
    let x = 2 in
      fun y -> (x :: y :: one)) : int \rightarrow int list
```

Which rule?

{one : int list} |- (1 :: one) : int list

Application

Constants Rule

Constants Rule

```
{one : int list} |- {one : int list} |- (::) : int \rightarrow int list \rightarrow int list
```

Rule for variables

{one: int list} |- one:int list

10/3/17 45

?

```
{x:int; one : int list} |- fun y -> (x :: y :: one))
: int \rightarrow int list
```

?

```
{y:int; x:int; one : int list} |-(x :: y :: one) : int list
{x:int; one : int list} |-fun y -> (x :: y :: one)
: int \rightarrow int list
```

```
6
```

7

```
{y:int; x:int; one:int list} {y:int; x:int; one:int list} |-((::) x):int list \rightarrow int list |-(y::one):int list} {y:int; x:int; one: int list} |-(x::y::one):int list {x:int; one: int list} |-(x::y::one):int list : int \rightarrow int list
```

Constant

Variable

```
{...} |- (::)

: int→ int list→ int list {...; x:int;...} |- x:int

{y:int; x:int; one : int list} |- ((::) x)

:int list→ int list
```



Curry - Howard Isomorphism

- Type Systems are logics; logics are type systems
- Types are propositions; propositions are types
- Terms are proofs; proofs are terms
- Function space arrow corresponds to implication; application corresponds to modus ponens



Curry - Howard Isomorphism

Modus Ponens

$$\begin{array}{c} A \Rightarrow B & A \\ \hline B & \end{array}$$

Application

$$\Gamma \mid -e_1 : \alpha \rightarrow \beta \quad \Gamma \mid -e_2 : \alpha$$

$$\Gamma \mid -(e_1 e_2) : \beta$$

Mea Culpa

- The above system can't handle polymorphism as in OCAML
- No type variables in type language (only metavariable in the logic)
- Would need:
 - Object level type variables and some kind of type quantification
 - let and let rec rules to introduce polymorphism
 - Explicit rule to eliminate (instantiate) polymorphism

-

Support for Polymorphic Types

- Monomorpic Types (τ):
 - Basic Types: int, bool, float, string, unit, ...
 - Type Variables: α , β , γ , δ , ϵ
 - Compound Types: $\alpha \rightarrow \beta$, int * string, bool list, ...
- Polymorphic Types:
 - Monomorphic types τ
 - Universally quantified monomorphic types
 - \blacksquare $\forall \alpha_1, \ldots, \alpha_n . \tau$
 - Can think of τ as same as $\forall \cdot \tau$

Support for Polymorphic Types

- Typing Environment \(\Gamma\) supplies polymorphic types (which will often just be monomorphic) for variables
- Free variables of monomorphic type just type variables that occur in it
 - Write FreeVars(τ)
- Free variables of polymorphic type removes variables that are universally quantified
 - FreeVars($\forall \alpha_1, \dots, \alpha_n \cdot \tau$) = FreeVars(τ) { $\alpha_1, \dots, \alpha_n$ }
- FreeVars(Γ) = all FreeVars of types in range of Γ

Monomorphic to Polymorphic

- Given:
 - type environment
 - monomorphic type τ
 - T shares type variables with I
- Want most polymorphic type for that doesn't break sharing type variables with
- Gen $(\tau, \Gamma) = \forall \alpha_1, ..., \alpha_n \cdot \tau$ where $\{\alpha_1, ..., \alpha_n\} = \text{freeVars}(\tau) \text{freeVars}(\Gamma)$

Polymorphic Typing Rules

A type judgement has the form

$$\Gamma$$
 |- exp : τ

- I uses polymorphic types
- τ still monomorphic
- Most rules stay same (except use more general typing environments)
- Rules that change:
 - Variables
 - Let and Let Rec
 - Allow polymorphic constants
- Worth noting functions again



Polymorphic Let and Let Rec

let rule:

$$\Gamma \mid -e_1 : \tau_1 \{x : Gen(\tau_1, \Gamma)\} + \Gamma \mid -e_2 : \tau_2 \}$$

$$\Gamma \mid -(let x = e_1 in e_2) : \tau_2$$

let rec rule:

$$\{x : \tau_1\} + \Gamma \mid -e_1:\tau_1 \{x: Gen(\tau_1, \Gamma)\} + \Gamma \mid -e_2:\tau_2 \}$$

$$\Gamma \mid -(let rec x = e_1 in e_2) : \tau_2$$

Polymorphic Variables (Identifiers)

Variable axiom:

$$\Gamma \mid -x : \varphi(\tau)$$
 if $\Gamma(x) = \forall \alpha_1, ..., \alpha_n . \tau$

- Where φ replaces all occurrences of $\alpha_1, \ldots, \alpha_n$ by monotypes τ_1, \ldots, τ_n
- Note: Monomorphic rule special case:

$$\Gamma \mid -x : \tau$$
 if $\Gamma(x) = \tau$

Constants treated same way



Fun Rule Stays the Same

fun rule:

$$\{x : \tau_1\} + \Gamma \mid -e : \tau_2$$

$$\Gamma \mid -\text{ fun } x -> e : \tau_1 \to \tau_2$$

- Types τ_1 , τ_2 monomorphic
- Function argument must always be used at same type in function body



Polymorphic Example

- Assume additional constants:
- hd : $\forall \alpha$. α list -> α
- tl: $\forall \alpha$. α list -> α list
- is_empty : $\forall \alpha$. α list -> bool
- \blacksquare :: : $\forall \alpha$. α -> α list -> α list
- \blacksquare [] : $\forall \alpha$. α list

Polymorphic Example

Show:

?

```
{} |- let rec length =
    fun | -> if is_empty | then 0
        else 1 + length (tl | l)
    in length ((::) 2 []) + length((::) true []) : int
```

Polymorphic Example: Let Rec Rule

```
• Show: (1)
                                    (2)
{length: \alpha list -> int} {length: \forall \alpha. \alpha list -> int}
|- fun | -> ...
                           |- length ((::) 2 []) +
                              length((::) true []) : int
 : \alpha list -> int
{} |- let rec length =
       fun I -> if is_empty I then 0
                  else 1 + length (tl I)
 in length ((::) 2 []) + length((::) true []) : int
```



Polymorphic Example (1)

Show:

?

: α list -> int

Polymorphic Example (1): Fun Rule

```
• Show: (3)
{length:\alpha list -> int, |\alpha| list } |-
if is empty I then 0
    else length (hd l) + length (tl l) : int
\{length: \alpha list -> int\} \mid -
fun I -> if is_empty I then 0
                  else 1 + length (tl l)
: \alpha list -> int
```

Polymorphic Example (3)

- Let $\Gamma = \{ length : \alpha list -> int, l: \alpha list \}$
- Show

?

 Γ |- if is_empty | then 0 else 1 + length (tl | l) : int

Polymorphic Example (3):IfThenElse

- Let $\Gamma = \{ length : \alpha list -> int, l: \alpha list \}$
- Show

```
(4) (5) (6) \Gamma|-\text{ is\_empty } \Gamma|-\text{ 0:int } \Gamma|-\text{ 1} + \\ \text{: bool} \qquad \qquad \text{length (tl l) : int}
```

 Γ |- if is_empty | then 0 else 1 + length (tl | l) : int

•

Polymorphic Example (4)

- Let $\Gamma = \{ length : \alpha list -> int, l: \alpha list \}$
- Show

?

 Γ |- is_empty | : bool

•

Polymorphic Example (4):Application

- Let $\Gamma = \{ length : \alpha list -> int, l: \alpha list \}$
- Show

?

?

$$\Gamma$$
 - is_empty : α list -> bool

 Γ |-|: α list

 Γ |- is_empty | : bool

Polymorphic Example (4)

- Let $\Gamma = \{ length : \alpha list -> int, l: \alpha list \}$
- Show

```
By Const since \alpha list -> bool is instance of \forall \alpha. \alpha list -> bool ?
```

```
\Gamma | - is_empty : \alpha list -> bool \Gamma | - I : \alpha list
```

 Γ |- is_empty | : bool

Polymorphic Example (4)

- Let $\Gamma = \{ length : \alpha list -> int, l: \alpha list \}$
- Show

By Const since α list -> bool is By Variable instance of $\forall \alpha$. α list -> bool $\Gamma(I) = \alpha$ list

 Γ |- is_empty : α list -> bool Γ |- l : α list

 Γ |- is_empty | : bool

This finishes (4)



Polymorphic Example (5):Const

- Let $\Gamma = \{ length : \alpha list -> int, l: \alpha list \}$
- Show

By Const Rule

Γ|- 0:int

Polymorphic Example (6):Arith Op

- Let $\Gamma = \{ length : \alpha list -> int, l: \alpha list \}$
- Show

 Γ |-1 + length (tl l) : int

Polymorphic Example (7):App Rule

- Let $\Gamma = \{ length : \alpha list -> int, l: \alpha list \}$
- Show

$$\Gamma$$
 |- (tl l) : α list -> α list

By Variable

$$\Gamma$$
|-|: α list

 Γ - (tl l) : α list

By Const since α list -> α list is instance of $\forall \alpha$. α list -> α list

Polymorphic Example: (2) by ArithOp

- Let $\Gamma' = \{ \text{length} \not\exists \alpha. \alpha \text{ list -> int} \}$
- Show:

```
(8) (9) \Gamma' |- \Gamma' |- \Gamma' |- length ((::) 2 []) :int length((::) true []) : int {length: \alpha. \alpha list -> int} |- length ((::) 2 []) + length((::) true []) : int
```

-

Polymorphic Example: (8)AppRule

- Let $\Gamma' = \{ \text{length} \not\exists \alpha. \alpha \text{ list -> int} \}$
- Show:

$$\Gamma'$$
 |- length : int list ->int Γ' |- ((::)2 []):int list

$$\Gamma'$$
 |- length ((::) 2 []) :int

Polymorphic Example: (8)AppRule

- Let $\Gamma' = \{ \text{length} \not\exists \alpha. \alpha \text{ list -> int} \}$
- Show:

By Var since int list -> int is instance of $\forall \alpha$. α list -> int

```
\frac{\Gamma' \mid - \text{ length : int list ->int } \Gamma' \mid - ((::)2 \mid ]):\text{int}}{\text{list}}
```

 Γ' |- length ((::) 2 []) :int

Polymorphic Example: (10)AppRule

- Let $\Gamma' = \{ \text{length} \not\exists \alpha. \alpha \text{ list -> int} \}$
- Show:
- By Const since α list is instance of $\forall \alpha$. α list

(11) $\Gamma' \mid -((::) 2) : \text{ int list } -> \text{ int list } \Gamma' \mid -[] : \text{ int list } \Gamma' \mid -((::) 2 \]] : \text{ int list }$

Polymorphic Example: (11)AppRule

- Let $\Gamma' = \{ \text{length} \not\exists \alpha. \alpha \text{ list -> int} \}$
- Show:
- By Const since α list is instance of

 $\forall \alpha$. α list

By Const

$$\Gamma'$$
 |- (::) : int -> int list -> int list int

$$\Gamma'$$
 |- ((::) 2) : int list -> int list

Polymorphic Example: (9)AppRule

- Let $\Gamma' = \{ \text{length} \not\exists \alpha. \alpha \text{ list -> int} \}$
- Show:

```
\Gamma' |- \Gamma' |- length:bool list ->int (::) true []:bool list \Gamma' |- length (::) true [] :int
```

Polymorphic Example: (9)AppRule

- Let $\Gamma' = \{ \text{length} \not\exists \alpha. \alpha \text{ list -> int} \}$
- Show:

By Var since bool list -> int is instance of $\forall \alpha$. α list -> int

```
\Gamma' |- \Gamma' |- length:bool list ->int \Gamma' |- length (::) true []):bool list \Gamma' |- length ((::) true []) :int
```

-

Polymorphic Example: (12)AppRule

- Let $\Gamma' = \{ \text{length} \not\exists \alpha. \alpha \text{ list -> int} \}$
- Show:
- By Const since α list is instance of $\forall \alpha$. α list

(13)

 Γ' |-((::)true):bool list ->bool list Γ' |- []:bool list

 Γ' [- ((::) true []) :bool list

Polymorphic Example: (13)AppRule

- Let $\Gamma' = \{ \text{length} \not\exists \alpha. \alpha \text{ list -> int} \}$
- Show:

```
By Const since bool list is instance of \forall \alpha. \alpha list
```

By Const

```
\Gamma' |-
```

(::):bool ->bool list ->bool list true : bool

 Γ' |- ((::) true) : bool list -> bool list