Programming Languages and Techniques (CIS1200)

Lecture 11

Abstract types: Finite Maps Chapter 10

Announcements (1)

- Homework 3 is due tomorrow at 11.59pm
 - Practice with BSTs, generic functions, first-class functions, and abstract types
- Reading: Chapters 8, 9, and 10 of the lecture notes

Announcements (2)

- Midterm 1: Friday, September 27th
 - Coverage: up to Wednesday, Sep 25th (Chapters 1-10)
 - During lecture
 Last names: A Z
 Meyerson Hall B1
 - 60 minutes; closed book, closed notes
 - Review Material
 - old exams on the web site ("schedule" tab)
 - Review Session
 - Wednesday, September 25, 7:00-9:00pm, Towne 100 (will be recorded)
 - Review Videos available on canvas

Review: Abstract types (e.g., set)

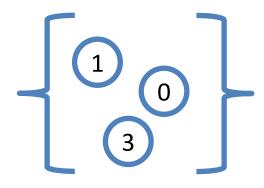
- An abstract type is defined by its *interface* and its *properties*, not its representation.
- Interface: defines operations on the type
 - There is an empty set
 - There is a way to add elements to a set to make a bigger set
 - There is a way to list all elements in a set
 - There is a way to test membership
- Properties: define how the operations interact with each other
 - Elements that were added can be found in the set
 - Adding an element a second time doesn't change the elements of a set
 - Adding in a different order doesn't change the elements of a set
- Any type (possibly with invariants) that satisfies the interface and properties can be a set
- Clients of an implementation can only access what is explicitly mentioned in the abstract type's interface



concrete representation

--- Interface ---

abstract view



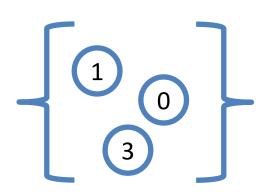
Review: Abstract vs. Concrete ULSet

```
module ULSet : SET = struct
  type 'a set = 'a list
  let empty : 'a set = []
  let add (x:'a) (s:'a set) :'a set =
      (* can treat s as a list *)
      x :: s
end
```

```
abstract view
```

= 0::3::1::[7]

S



```
module type SET = sig
  type 'a set
  val empty : 'a set
  val add : 'a -> 'a set -> 'a set
end
(* A client of the module *)
;; open ULSet
let s : int set
  = add 0 (add 3 (add 1 empty))
Client code doesn't (can't!) care
  about internal representation!
```

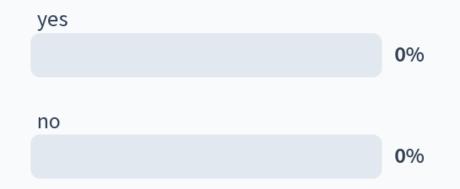
Review: Abstract vs. Concrete OLSet

module OLSet : SET = struct

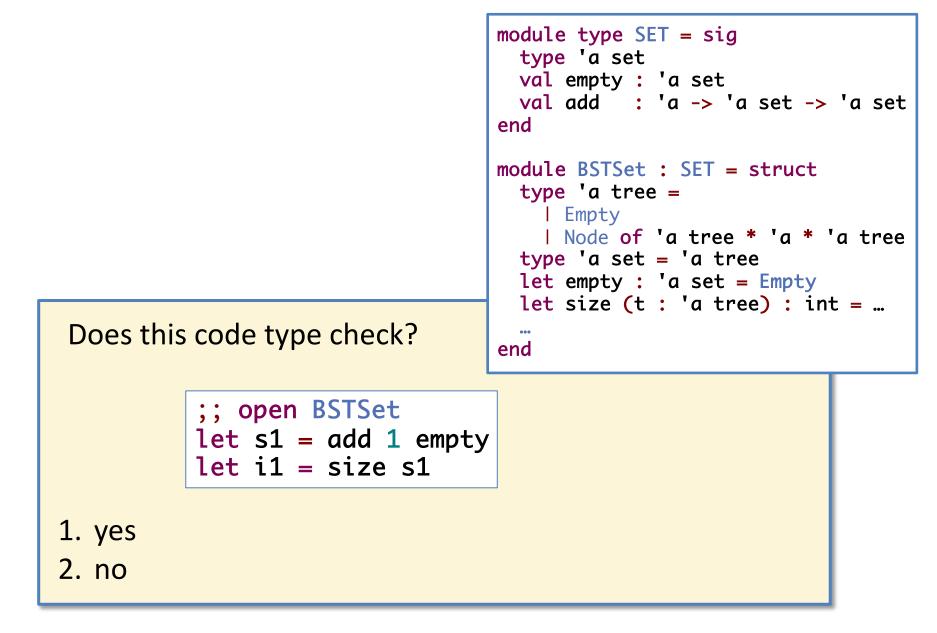
```
type 'a set = 'a list
                                 let empty : 'a set = []
                                 let add (x:'a) (s:'a set) :'a set =
                                   (* can treat s as a list, but
                                      must find right place for x *)
s = 0::1::3::[7]
                              end
                                  module type SET = sig
 concrete representation
                                    type 'a set
                                    val empty : 'a set
      abstract view
                                    val add : 'a \rightarrow 'a set \rightarrow 'a set
                                (* A client of the OLSet module *)
                                ;; open OLSet
                                let s : int set
                                  = add 0 (add 3 (add 1 empty))
                                                 Client code doesn't change!
```

10: Does this code typecheck?





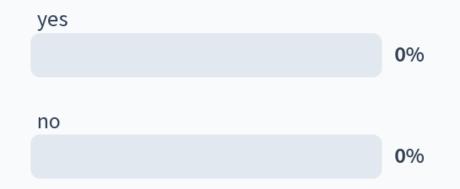
Start the presentation to see live content. For screen share software, share the entire screen. Get help at pollev.com/app



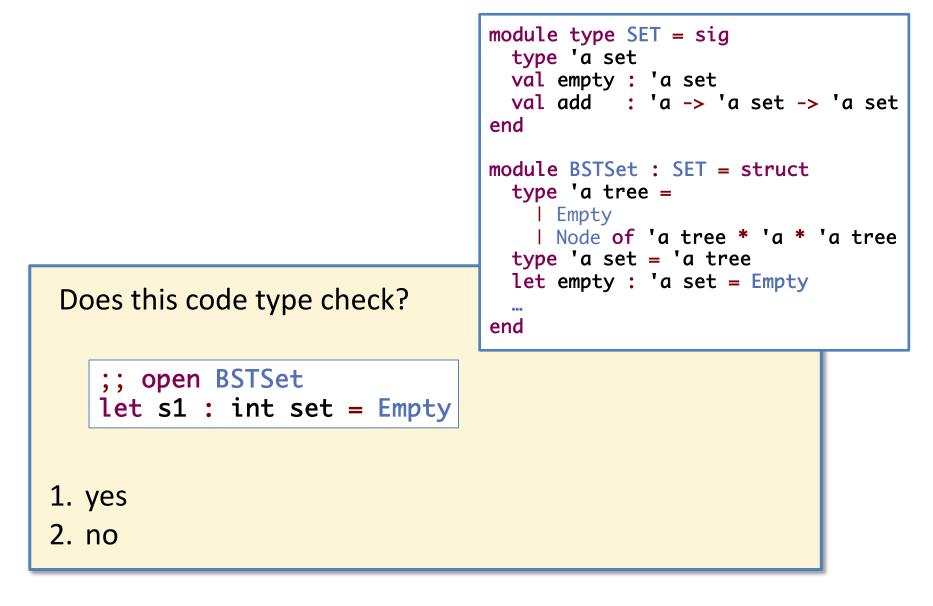
Answer: no, cannot access helper functions outside the module

10: Does this code typecheck?

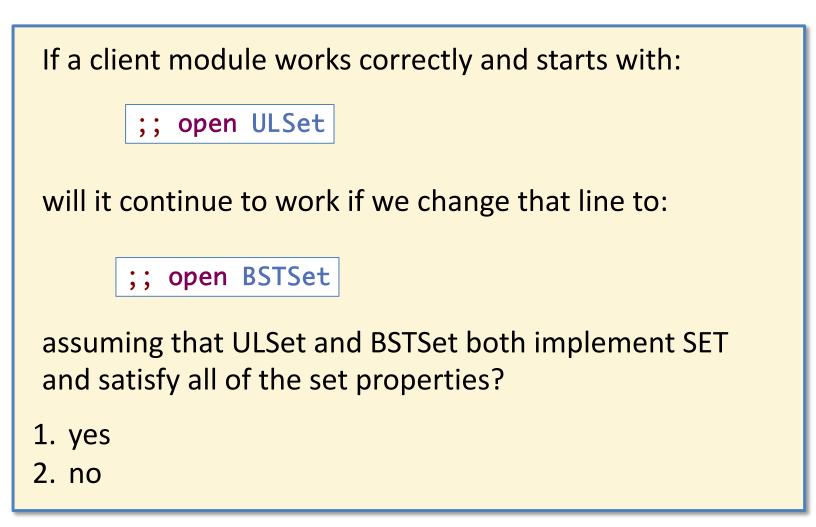




Start the presentation to see live content. For screen share software, share the entire screen. Get help at pollev.com/app



Answer: no, the Empty data constructor is not available outside the module



Answer: yes (though performance may be different)

Is it possible for a client to call **member** with a tree that is not a BST?

- 1. yes
- 2. no

No: the BSTSet operations preserve the BST invariants. there is no way to construct a non-BST tree using the interface.

Equality of Sets

When defining an abstract type, you may need to define a different notion of equality

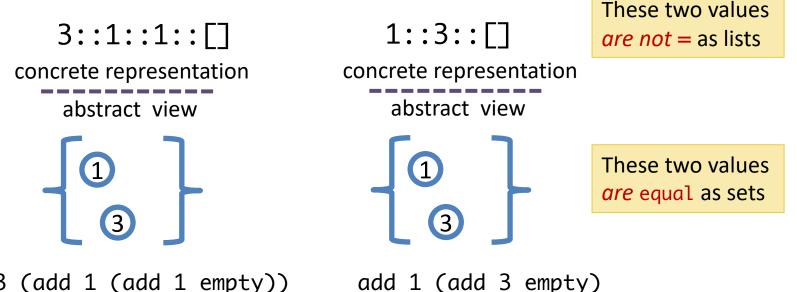
- The built-in "structural equality" (written =) may not be appropriate
- Be sure to use the 'equals' function when comparing, e.g., sets
- (Other generic operations, like < and > may also be affected.)

Equality of Sets

The SET interface includes

val equals : 'a set -> 'a set -> bool

- Can we use OCaml's built-in = to compare sets? •
 - This generic, built-in equality operation = compares the structure of its two inputs to see whether they are the same
- With unordered lists, NO!



true when both sets contain same elements

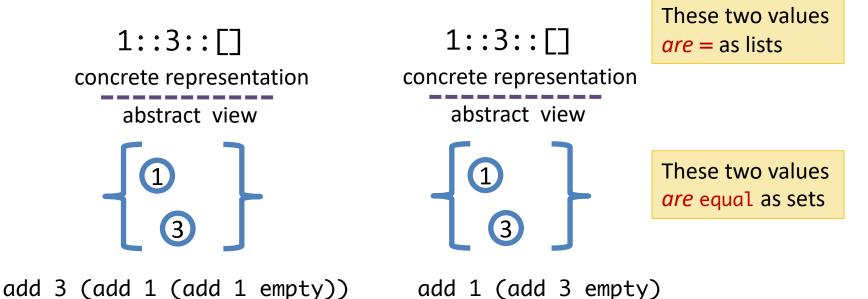
This function should return

Equality of Sets

• The SET interface includes

val equals : 'a set -> 'a set -> bool

- Can we use OCaml's built-in `=` to compare sets?
 - This generic, built-in equality operation = compares the *structure* of its two inputs to see whether they are the same
- With strictly ordered lists, YES!



This function should return

true when both sets

contain same elements

Abstract types: **BIG IDEA**

Hide the *concrete representation* of a type behind an *abstract interface* to preserve **representation invariants**

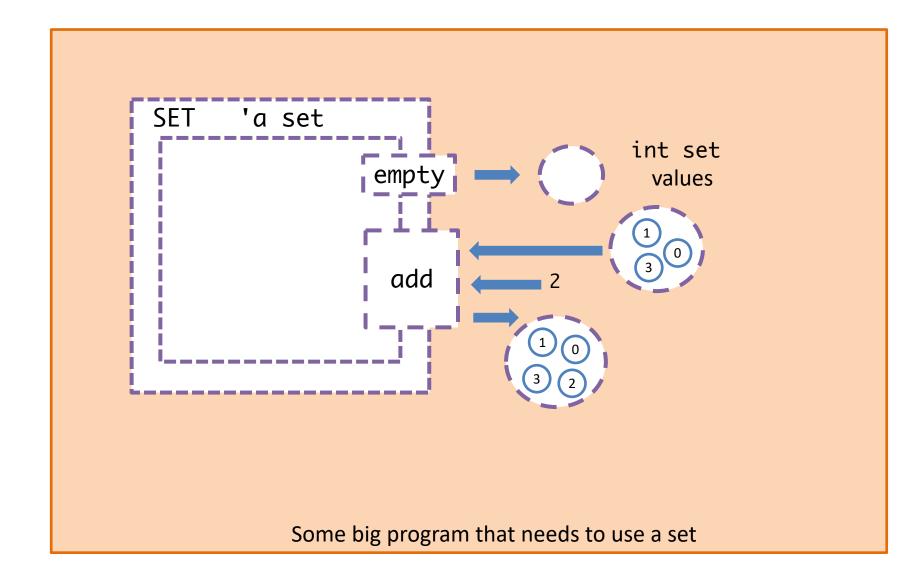
- Example representation invariants
 - Sets implemented as lists, which must be strictly ordered (no duplicates)
 - Sets implemented as binary tree, satisfying the BST invariant
- If the set type is abstract, and *all* operations preserve invariants, then invariants **must** hold for *all* sets in the program!
 - Example: if all sets implemented as lists are strictly ordered, then the `=` operation implements set equality
 - Example: if all sets implemented as trees satisfy the BST invariant, then the lookup function can *assume* that its input is a BST

Abstract types: **BIG IDEA**

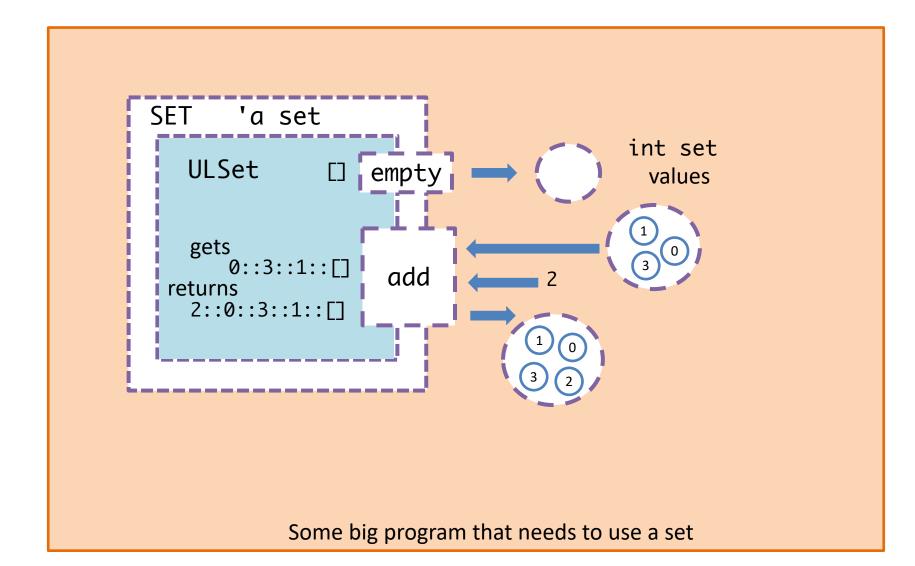
Hide the *concrete representation* of a type behind an *abstract interface* to preserve **representation invariants**

- An abstract interface **restricts** how other parts of the program can interact with the data
 - Type checking ensures that the **only** way to create a set is with the operations in the interface (empty, add, etc.)
 - Type checking ensures that clients cannot depend on whether the sets are implemented as trees or lists
- Benefits
 - Safety: The other parts of the program can't violate invariants, which would cause bugs
 - Modularity: It is possible to change the implementation without changing the rest of the program

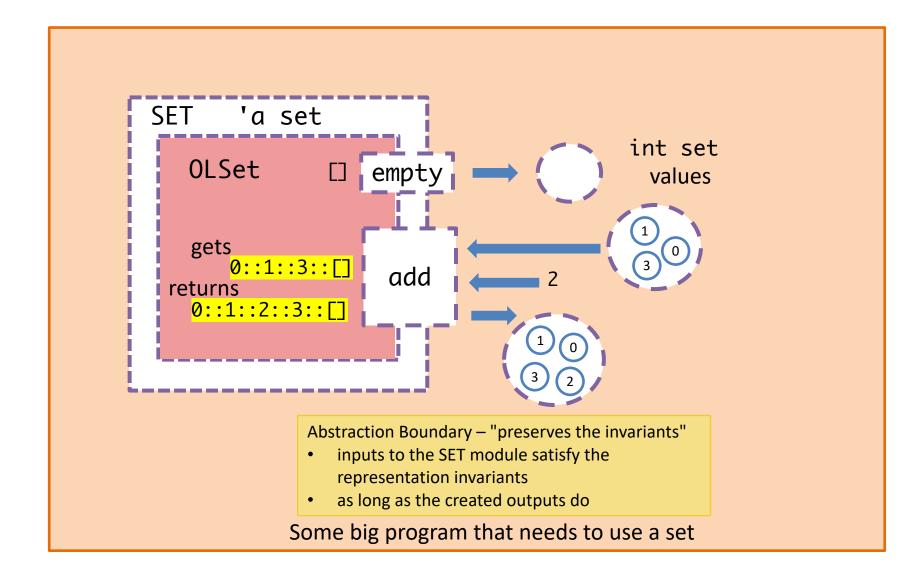
Encapsulation and Modularity



Implementation



Implementation



Property-Based Testing

Testing Styles

- "From the inside"...
 - If we know the concrete representation of our data, we can test the effect of each operation on that representation
 - Useful for checking that invariants are maintained
- "From the outside"...
 - If the concrete representation is hidden, this doesn't work!
 - We need a different way to think about testing

What Should We Test?

- Interface: Names and types of operations on the abstract type
- Properties: How the operations behave and interact
 - "Elements that were added can be found by lookup"
 - "Adding an element a second time doesn't change the elements of a set
 - "Adding elements in a different order doesn't change the outcome of later operations"

Test the properties!

A *property* is a general statement about the behavior of functions in the interface. E.g.,

For any set s and any element x, member x (add x s) = true

A good test case *checks a specific instance* of the property: let test () : bool = (member 3 (add 3 empty))
;; run_test "member 3 (add 3 empty)" test

Property-based Testing

1. Translate informal requirements into general statements about the interface.

```
Example: "Order doesn't matter" becomes
For any set s and any elements x and y,
add x (add y s) "equals" add y (add x s)
```

2. Write tests for the "interesting" instances of the general statement.

Example "interesting" choices:

- s is empty vs. s is nonempty
- x = y vs. x <> y
- x and/or y already in s
 vs. x and y different from what's in s

Notes:

- Not usually possible to exhaustively test all possibilities (too many!): so just try to cover the "interesting" choices
- Be careful with equality! ULSet.equals is *not* the same as =.

11: How comfortable are you with abstract types, interfaces, and implementations?			
	Totally Lost	0%	
	Still working on understanding		
		0%	
	I think I mostly get it		
		0%	
	Pretty comfortable - it makes sense		
		0%	
	Ready to design my own ADTs		
		0%	

Finite Maps

A case study on abstract interfaces and concrete implementations

Motivating Scenario

- Suppose you were writing a course-management system and needed to look up the lab section for a student given the student's PennKey...
 - Students might add/drop the course
 - Students might switch lab sections
 - Students should be in only one lab section
- How would you do it? What data structure would you use?

Key/Value store

Кеу	Value
"stephanie"	15
"mitch"	05
"ezaan"	10
"likat"	15

- Each key is associated with a value.
 - No two keys are identical
 - Values can be repeated
- Given the key "stephanie", we want to find / lookup the value 15

Finite Maps

- A *finite map* (a.k.a. *dictionary*) is a collection of *entries* from distinct *keys* to *values*.
 - Operations to *add* a new entry, *test* for key membership, *get* the value bound to a particular key, *list* all entries stored in the map
- Example: we might use a finite map to look up the lab section of a CIS 1200 student
- Like sets, *finite maps* appear in many settings:
 - domain names to IP addresses
 - words
 to their definitions (a dictionary)
 - user names
 to passwords

Signature: Finite Map

Design Process Step 2: specify the interface

<pre>module type MAP = sig</pre>					
		AI = 319	The map type is generic in <i>two</i> ways:		
	type ('k,'v) map	type of keys and type of values		
	val empty	: ('k,'v) m	ap		
	val add	: 'k -> 'v	-> ('k,'v) map -> ('k,'v) map		
	val mem	: 'k -> ('k	,'v) map -> bool		
	val get	: 'k -> ('k	(,'v) map -> 'v		
	<mark>val</mark> equals	: ('k,'v) m	ap -> ('k,'v) map -> bool		
	end				

Properties of Finite Maps

For any finite map m, key k, and value v:

- 1. get k (add k v m) = v
- 2. If k1 <> k2 then get k1 (add k2 v2 (add k1 v1 m)) = v1
- 3. If mem k m = true then there is a v such that get k m = v

5. mem k (add k v m) = true

(among others...)

Design Process Step 3: write test cases

Tests for Finite Map abstract type

;; open Assert

Design Process Step 3: write test cases

(* Specifying the properties of the MAP abstract type via test cases. *)

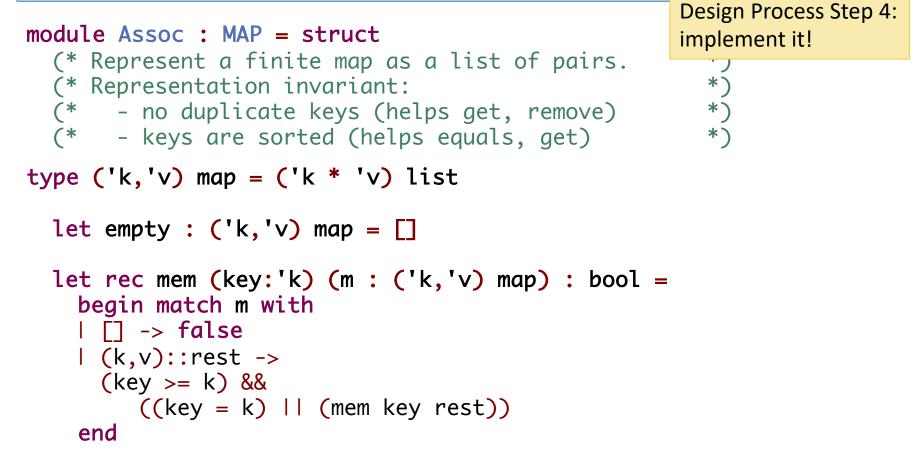
```
Using an anonymous
(* A simple map with one element. *)
                                                              function avoids making up a
let m1 : (int,string) map = add 1 "uno" empty
                                                              (redundant) function name
                                                              for the test
(* access value for key in the map *)
;; run_test "find 1 m1" (fun () -> (get 1 m1) = "uno")
(* find for value that does not exist in the map? *)
;; run failing test "find 2 m1" (fun () -> (get 2 m1) = "dos" )
let m2 : (int, string) map = add 1 "un" m1
(* find after redefining value, should be new value *)
;; run_test "find 1 m2" (fun () -> (get 1 m2) = "un")
(* test membership *)
;; run test "mem test" (fun () ->
        mem 1 (add 2 "dos" (add 1 "uno" empty)))
```

Finite Map Demo

Implementing the module

finiteMap.ml

Implementation: Ordered Lists



;; run_test "mem test" (fun () -> mem "b" [("a",3); ("b",4)])

Implementation: Ordered Lists

```
let rec get (key:'k) (m : ('k,'v) map) : 'v =
  begin match m with
  | [] -> failwith "key not found"
  | (k,v)::rest ->
    if key < k then failwith "key not found"
    else if key = k then v
   else get key rest
  end
let rec remove (key:'k) (m : ('k,'v) map) : ('k,'v) map =
  begin match m with
  | [] -> []
  | (k,v)::rest ->
    if key < k then m
    else if key = k then rest
   else (k,v)::remove key rest
  end
```

Summary: Abstract Types

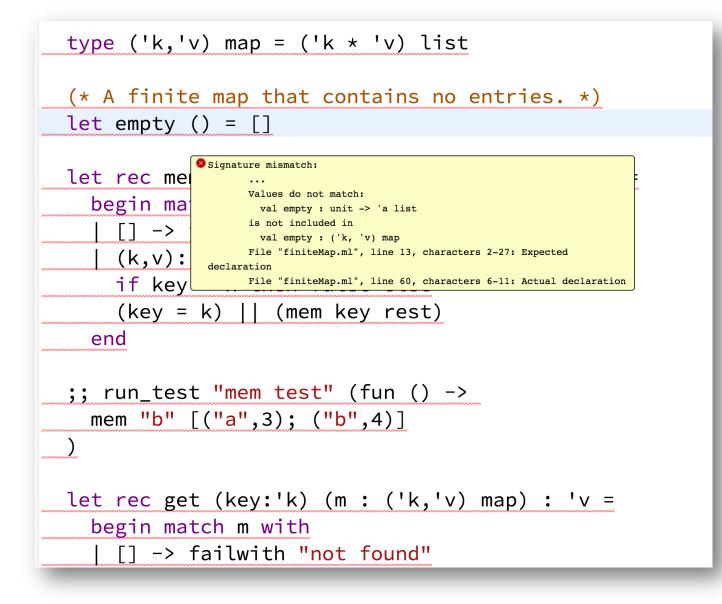
- Different programming languages support different ways of defining abstract types
- At a minimum, this means providing:
 - A way to specify (write down) an interface
 - A means of hiding implementation details (encapsulation)
- In OCaml:
 - Interfaces are specified using a *signature* or *interface*
 - Encapsulation: the interface can *omit* information
 - type definitions
 - names of auxiliary functions
 - Clients *cannot* mention values or types not named in the interface

Typechecking

How does OCaml* typecheck your code?

*Historical aside: the algorithm we are about to see is known as the Damas-Hindley-Milner type inference algorithm. Turing Award winner Robin Milner was, among other things, the inventor of "ML" (for "meta language"), from which OCaml gets its "ml".

OCaml Typechecking Errors



Typechecking

How do we determine the type of an expression?

- 1. Recursively determine the types of *all* sub-expressions
 - Constants have "obvious" types
 - 3 : int "foo" : string true : bool
 - Identifiers may have type annotations
 - let and function arguments
 - Module signatures/interfaces
- 2. Expressions that *construct* structured values have compound types built from the types of sub-expressions

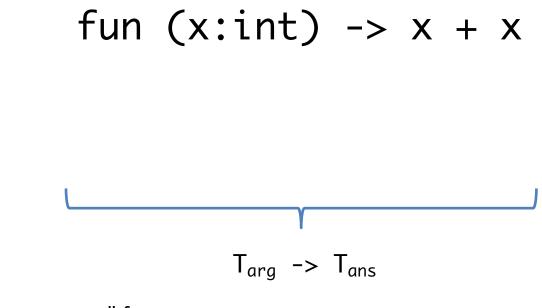
(3, "foo") : int * string
(fun (x:int) -> x + 1) : int -> int
Node(Empty, (3, "foo"), Empty) : (int * string) tree

To typecheck a function:

fun (x:int) -> x + x

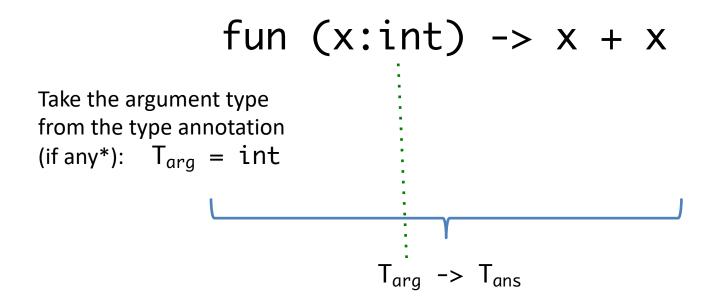


To typecheck a function:



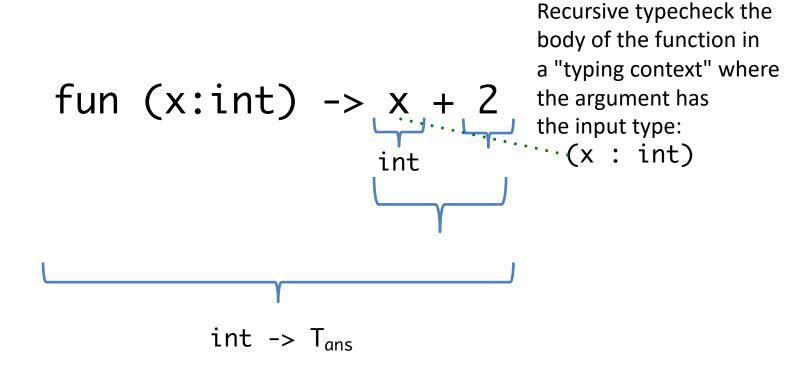
Make up "new names" for the input (argument) and output (answer) types.

To typecheck a function:

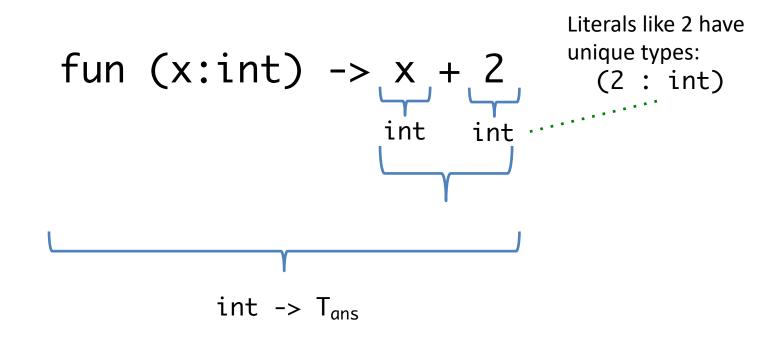


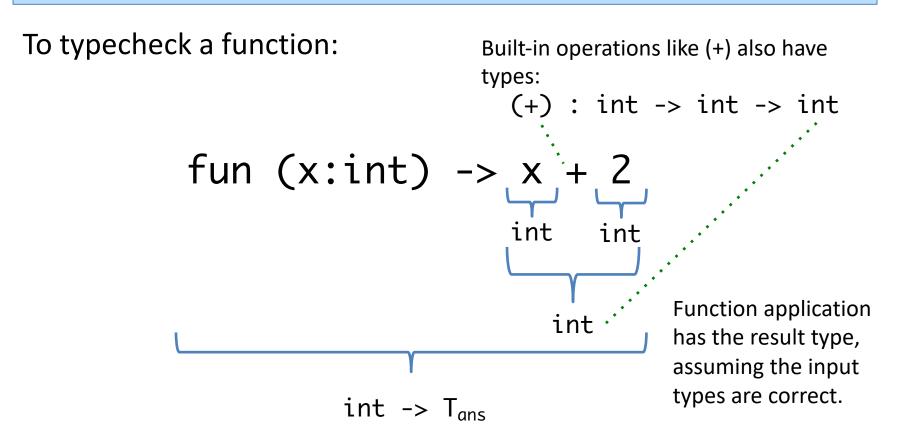
*If there is no annotation, just use the "fresh" name...

To typecheck a function:

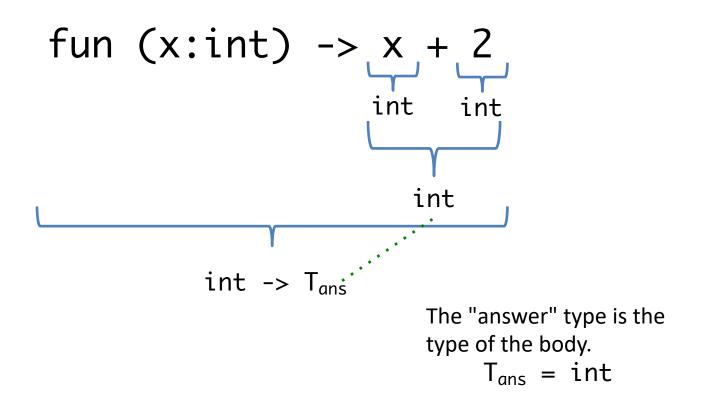


To typecheck a function:

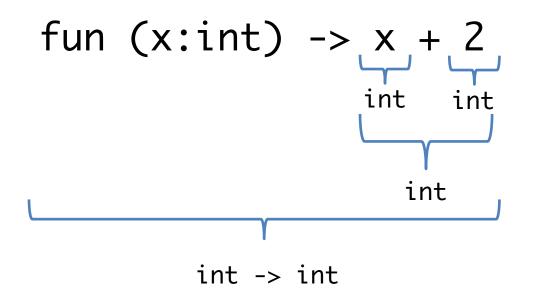




To typecheck a function:



To typecheck a function:



- 3. The type of a function-application expression is obtained as the result from the function type:
 - Given a function f : T_{arg} -> T_{ans}
 - and an argument e : T_{arg}
 - the application (f e) : T_{ans}

has the answer type

((fun (x:int) (y:bool) -> y) 3) : ??

- 3. The type of a function-application expression is obtained as the result from the function type:
 - Given a function f : T_{arg} -> T_{ans}
 - and an argument e : T_{arg}
 - the application (f e) : T_{ans}

has the answer type

((fun (x:int) (y:bool) -> y) 3) : ??

- 3. The type of a function-application expression is obtained as the result from the function type:
 - Given a function f : T_{arg} -> T_{ans}
 - and an argument
 e
 : T
 - the application (f e) : T_{ans}

- 3. The type of a function-application expression is obtained as the result from the function type:
 - f Given a function : T_a
 - and an argument e : T_{arg} the application (f e) : T_{ans}

- 3. The type of a function-application expression is obtained as the result from the function type:
 - Given a function f : T_{arg} -> T_{ans}
 - and an argument
 e
 - the application (fe):

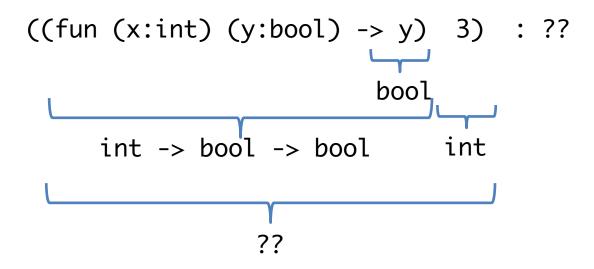
$$T_{arg} \rightarrow T_{ans}$$

 T_{arg} of the input type
 T_{ans} has the answer type

- 3. The type of a function-application expression is obtained as the result from the function type:
 - Given a function f : T_{arg} -> T_{ans}
 - and an argument
 e
 - the application (f e) : T_{ans}

$$T_{arg} \rightarrow T_{ans}$$

 T_{arg} of the input type
 T_{arg} has the answer type



- 3. The type of a function-application expression is obtained as the result from the function type:
 - Given a function f : $T_{arg} \rightarrow T_{ans}$
 - and an argumente: T_{arg} of the input type- the application(f e): T_{ans} has the answer type

 $((fun (x:int) (y:bool) \rightarrow y) 3) : ??$ $int \rightarrow bool \rightarrow bool \quad int$ $T_1 = int$ $T_2 = bool \rightarrow bool$

- What about generics? i.e., what if f:'a ->'a?
- For generic types we *unify*
 - Given a function f : $T_1 \rightarrow T_2$
 - and an argument e : U_1 of the input type Can *"match up"* T_1 and U_1 to obtain information about type parameters in T_1 and U_1 based on their usage
- Unification:
 - try to match up corresponding parts of the type
 (int list) tree ⇔ 'a tree

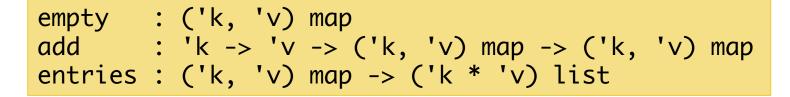
– Obtain an *instantiation*: e.g. 'a = int list

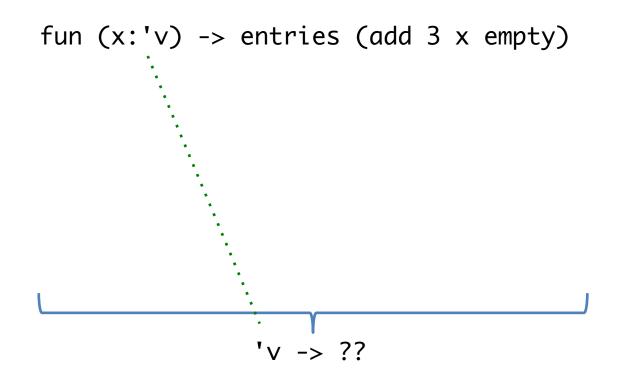
bool tree \Leftrightarrow

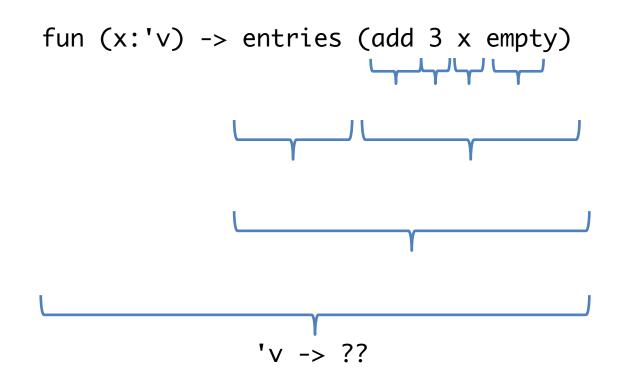
- Propagate that information to all occurrences of 'a
- If not possible, unification fails, meaning a type checking error

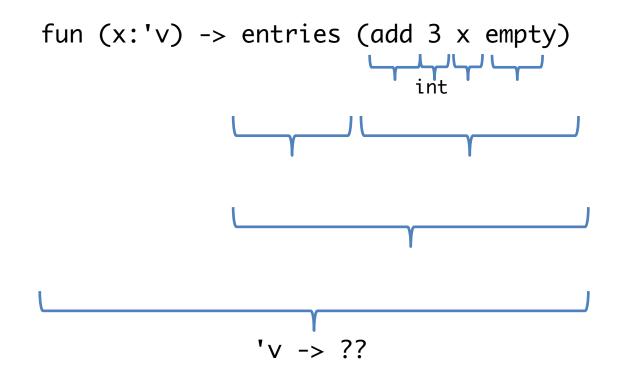
ERROR! bool \neq int

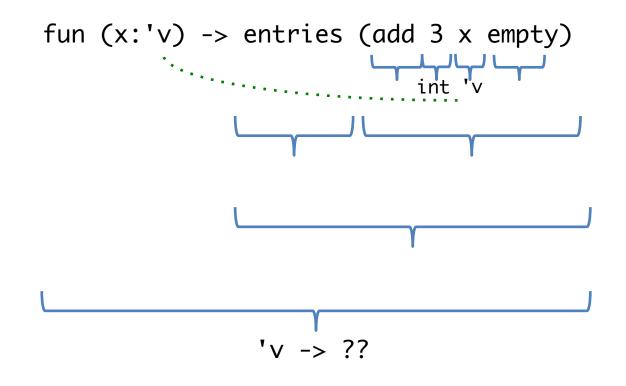
int tree

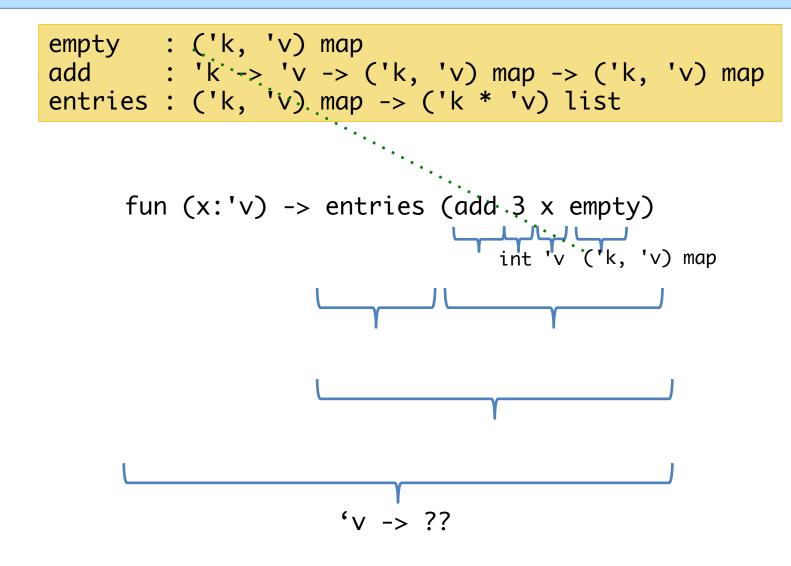


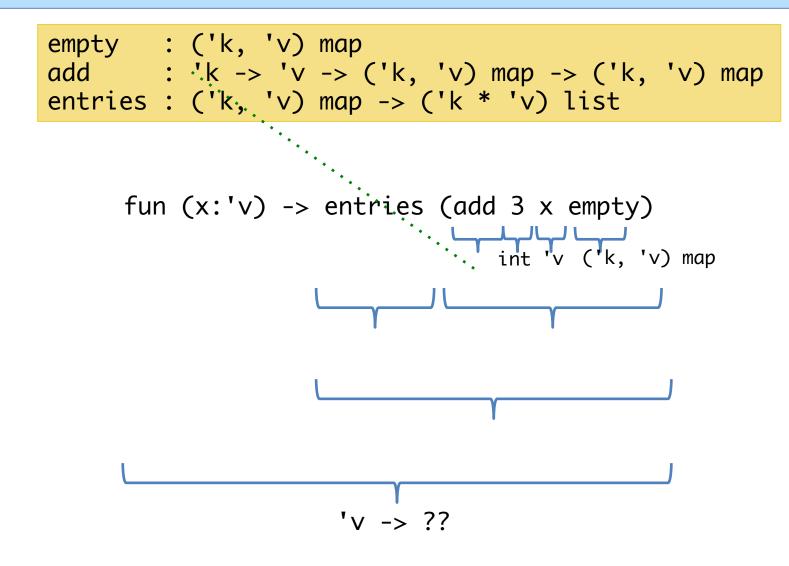


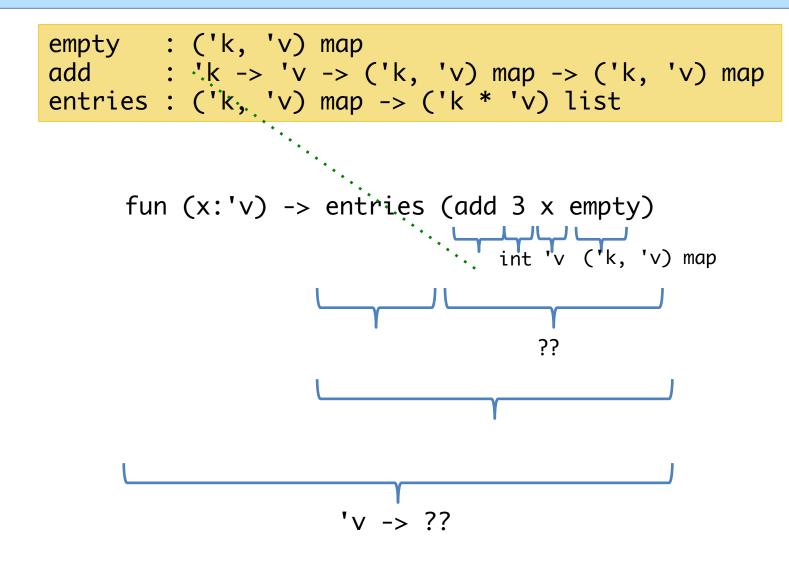


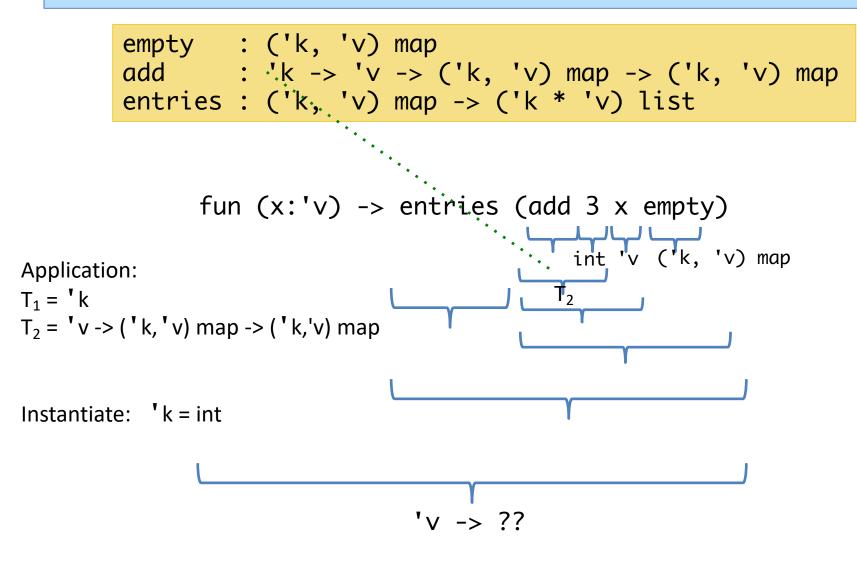


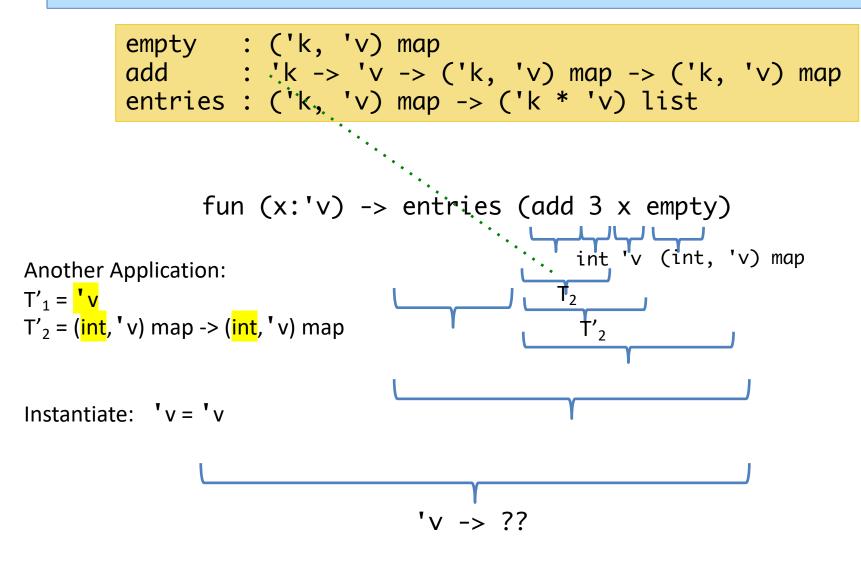


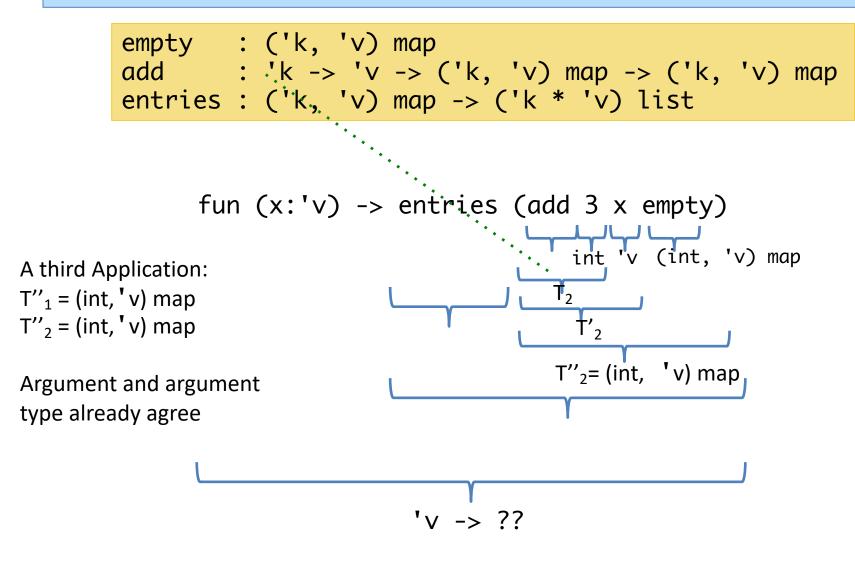


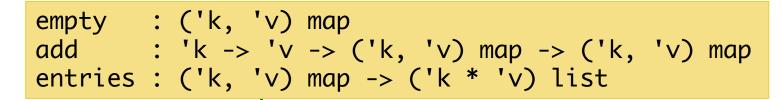


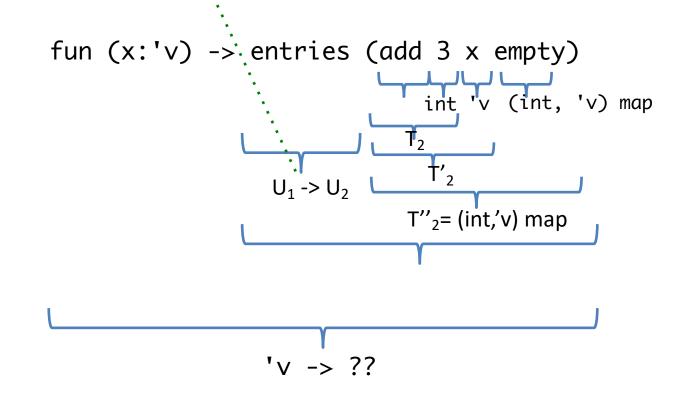


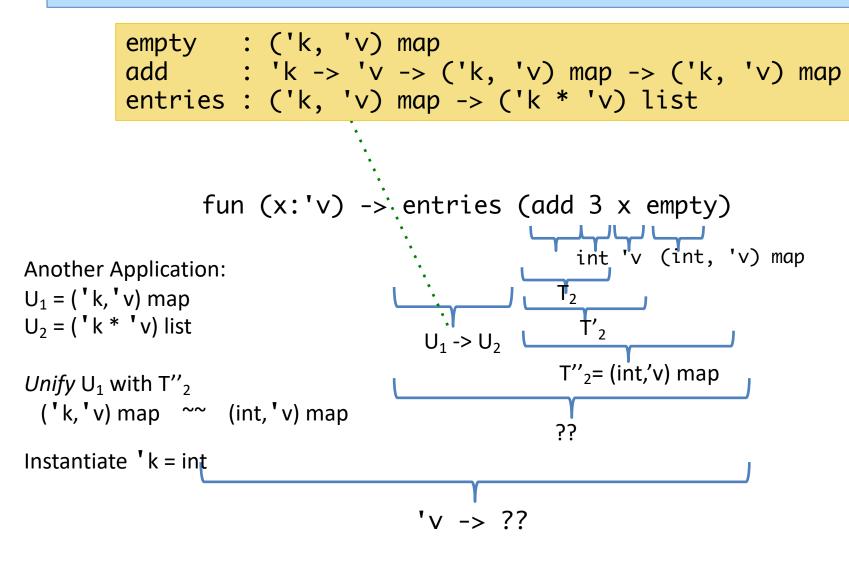


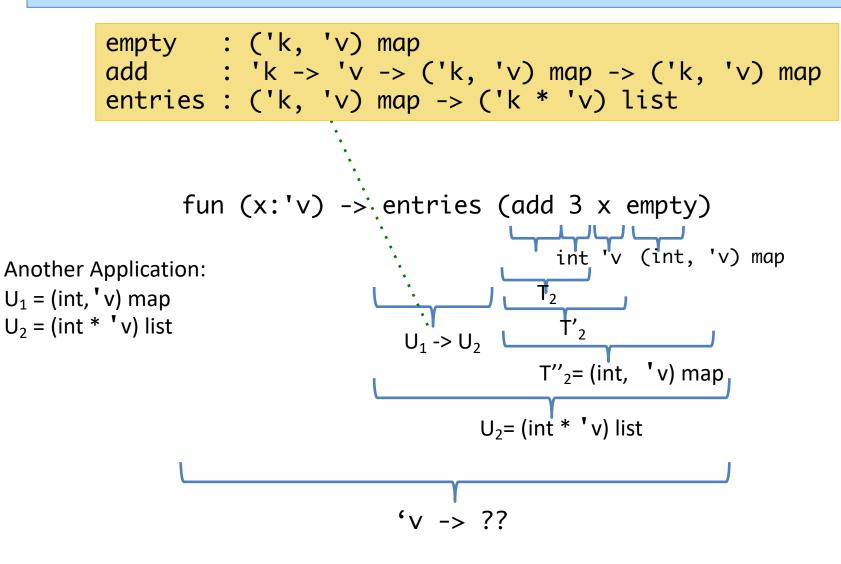


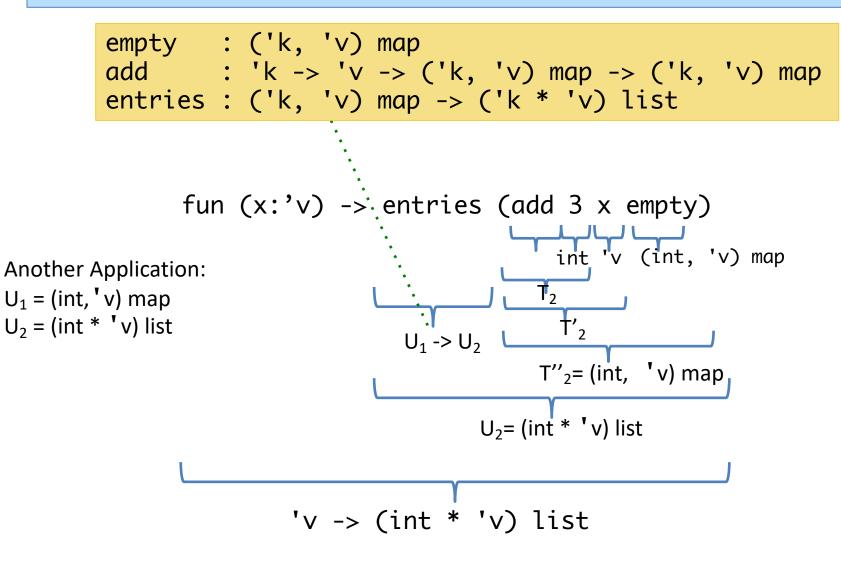












Ill-typed Expressions?

• An expression is ill-typed if, during this type checking process, inconsistent constraints are encountered:

add 3 true (add "foo" false empty)

Error: found int but expected string

12: What is the type of this expression?



