# Unions and Products

**Concepts of Programming Languages Lecture 7** 

#### Announcements

- » Thanks for providing feedback on the course!
- » Lectures will be more interactive: I will be asking more questions; we will do derivations and programming together
- » More examples! Each concept will be followed by an example on typing, semantic rules, and programming. More practice via homework.

#### Practice Problem 1 (Tail Recursive)

Implement the function
reverse : 'a list -> 'a list



#### Practice Problem 2 (Tail Recursive)

Implement the function double where double l doubles every element of the list l



#### Homework (Tail Recursive)

```
Implement the function
delete_every_other : 'a list -> 'a list
```



#### Outline

- » Discuss the use of tuples and records for creating collections of data
- » Introduce algebraic data types (ADTs) for creating data with given "shapes"
- » Cover parametric and recursive ADTs for more general data structures

## Products

```
let point : float * float = (2.0, 3.0)
let student : string * int = ("Franco", 244342)
```

```
let point : float * float = (2.0, 3.0)
let student : string * int = ("Franco", 244342)
```

Tuples are **ordered unlabeled** fixed—length heterogeneous collections of data

```
let point : float * float = (2.0, 3.0)
let student : string * int = ("Franco", 244342)
```

Tuples are **ordered unlabeled** fixed—length heterogeneous collections of data

(I expect that these are familiar)

```
let point : float * float = (2.0, 3.0)
let student : string * int = ("Franco", 244342)
```

Tuples are **ordered unlabeled** fixed—length heterogeneous collections of data

(I expect that these are familiar)

These are useful for returning multiple values from a function

#### Pattern Matching on Tuples

```
let dist (p1 : float * float) (p2 : float * float) : float =
  match p1, p2 with
  | (x1, y1), (x2, y2) ->
    let x = x1 -. x2 in
    let y = y1 -. y2 in
    sqrt ((x *. x) +. (y *. y))
```

There are no accessors for tuples

Instead we can use pattern matching

A pattern is a typed template for how a piece of data should look

A **pattern** is a typed template for how a piece of data should look

A **match-expression** is a way of *destructing* <u>any</u> piece of data

A pattern is a typed template for how a piece of data should look

A match-expression is a way of destructing any piece of data

We  $\mathit{match}$  on an expression e, and check if the value of e  $\mathit{matches}$  with the pattern p

#### Note: Patterns are not Expressions

Patterns are similar to expressions, but with some key differences

They can be wildcards, they can be variables, there's a lot of possibilities (<u>link</u>)

#### Advanced Pattern Matching

```
let dist ((x1, y1) : float * float) ((x2, y2) : float * float) : float =
  let x = x1 - x2 in
 let y = y1 - y2 in
  sqrt((x *. x) +. (y *. y))
let dist (p1 : float * float) (p2 : float * float) : float =
 let (x1, y1) = p1 in
  let (x2, y2) = p2 in
  let x = x1 - x2 in
 let y = y1 - y2 in
  sqrt((x *. x) +. (y *. y))
```

Pattern matching can also be done implicitly in let-expressions and function arguments!

# And we can do all this formally!

#### Tuples (Syntax Rule)

If  $e_1, ..., e_n$  are well-formed expressions, then

$$(e_1, e_n)$$

is a well-formed expression



#### Tuples (Syntax Rule)

```
<expr> ::= ( <expr> , ... , <expr> )
```

If  $e_1, ..., e_n$  are well-formed expressions, then

```
(e_1, e_n)
```

is a well-formed expression

### Tuple (Typing Rule)

If  $e_1,...,e_n$  are of type  $\tau_1,...,\tau_n$ , respectively, in the context  $\Gamma$  then

$$(e_1, e_n)$$

is of type  $\tau_1$  \* ... \*  $\tau_n$  in the context  $\Gamma$ 



### Tuple (Typing Rule)

$$\frac{\Gamma \vdash e_1 : \tau_1 \qquad \dots \qquad \Gamma \vdash e_n : \tau_n}{\Gamma \vdash (e_1, \dots, e_n) : \tau_1 * \dots * \tau_n} \text{ (tuple)}$$

If  $e_1,...,e_n$  are of type  $\tau_1,...,\tau_n$ , respectively, in the context  $\Gamma$  then

$$(e_1, e_n)$$

is of type  $\tau_1$  \* ... \*  $\tau_n$  in the context  $\Gamma$ 

#### Tuple (Semantic Rule)

```
If e_1, \dots, e_n evaluate to v_1, \dots, v_n, respectively, then  ( \ e_1 \ , \ \dots \ , \ e_n \ )
```

evaluates to  $(v_1, \dots, v_n)$ 



#### Tuple (Semantic Rule)

$$\frac{e_1 \Downarrow v_1 \qquad \dots \qquad e_n \Downarrow v_n}{(e_1, \dots, e_n) \Downarrow (v_1, \dots, v_n)} \text{ (tupleEval)}$$

```
If e_1, \dots, e_n evaluate to v_1, \dots, v_n, respectively, then  ( \ e_1 \ , \ \dots \ , \ e_n \ )
```

evaluates to  $(v_1, \dots, v_n)$ 

```
\frac{\{x: int\} \vdash x: int}{\{x: int\} \vdash x + x : int} \xrightarrow{\{intAdd\}} \frac{\{x: int\} \vdash true : bool(tuple)}{\{x: int\} \vdash (x + x, true) : int * bool}
```

```
\{x : int\} \vdash (x + x, true) : int * bool
```

```
\{x: int\} \vdash x + x: int\} \vdash \{x: int\} \vdash true: bool
\{x: int\} \vdash (x + x, true): int * bool
```

```
\frac{\{x: int\} \vdash x: int}{\{x: int\} \vdash x + x: int} \qquad \qquad \{x: int\} \vdash true: bool \\ \{x: int\} \vdash (x + x, true): int * bool}
```

```
\frac{\{x: int\} \vdash x: int}{\{x: int\} \vdash x: int} \xrightarrow{\{x: int\} \vdash x: int} \xrightarrow{\{add\}} \{x: int\} \vdash true: bool 
\{x: int\} \vdash (x+x, true): int * bool
```

```
 \frac{\{x: int\} \vdash x: int \quad x: in
```

$$\frac{(\text{intLitEval})}{21/2} = \frac{31/3}{31/3} = \frac{(\text{intAdd})}{(\text{intAdd})} = \frac{(\text{trueLitEval})}{(\text{true})} = \frac{(\text{true})}{(\text{true})} = \frac{(\text{true})}{($$

```
(2 + 3, true) \Downarrow (5, T)
```

## Example (Semantic Derivation)

## Example (Semantic Derivation)

## Example (Semantic Derivation)

$$\frac{2 \Downarrow 2}{2 \Downarrow 2} \frac{(\text{intEval})}{(\text{addEval})} \frac{(\text{intEval})}{(\text{trueEval})} \frac{2 + 3 \Downarrow 5}{(\text{tupleEval})} \frac{(\text{trueEval})}{(\text{tupleEval})}$$

# Tuple Match (Typing Rule)

If p is of type  $\tau_1 * ... * \tau_n$  in the context  $\Gamma$  and e is of type  $\tau$  in the context  $\Gamma$  along with  $x_i$  of type  $\tau_i$ , then

```
match p with | x_1 \rangle \dots \langle x_n \rangle = e
```

is of type  $\tau$  in the context  $\Gamma$ 



# Tuple Match (Typing Rule)

$$\frac{\Gamma \vdash p : \tau_1 * \ldots * \tau_n \qquad \Gamma, x_1 : \tau_1, \ldots, x_n : \tau_n \vdash e : \tau}{\Gamma \vdash \mathsf{match} \ p \ \mathsf{with} \mid x_1, \ldots, x_n \to e : \tau} \ (\mathsf{matchTuple})$$

If p is of type  $\tau_1 * ... * \tau_n$  in the context  $\Gamma$  and e is of type  $\tau$  in the context  $\Gamma$  along with  $x_i$  of type  $\tau_i$ , then

match 
$$p$$
 with  $| x_1 , ... , x_n \rightarrow e$ 

is of type  $\tau$  in the context  $\Gamma$ 

# Tuple Match (Semantic Rule)

If p evaluates to  $(v_1, ..., v_n)$  and e evaluates to v after substituting  $v_i$  for  $x_i$  ( $\forall i$ ), then

match 
$$p$$
 with  $|(x_1, ..., x_n)| \rightarrow e$ 

evaluates to v



# Tuple Match (Semantic Rule)

$$\frac{p \Downarrow (v_1,\ldots,v_n)}{(\mathsf{match}\; p\; \mathsf{with}\; |\; x_1,\ldots,x_n \to e) \Downarrow v} \; (\mathsf{match}\; Tuple\mathsf{Eval})$$

If p evaluates to  $(v_1, ..., v_n)$  and e evaluates to v after substituting  $v_i$  for  $x_i$  ( $\forall i$ ), then

match 
$$p$$
 with  $x_1$ ,...,  $x_n \rightarrow e$ 

evaluates to v

#### Practice Problem

 $\emptyset$   $\vdash$  fun x  $\rightarrow$  match x with | (a, b)  $\rightarrow$  a + b :  $\tau$ 

Determine the type  $\tau$  so that the above judgment is derivable from the rules below. Also give a derivation

$$\frac{\Gamma \vdash p : \tau_1 * \ldots * \tau_n \qquad \Gamma, x_1 : \tau_1, \ldots, x_n : \tau_n \vdash e : \tau}{\Gamma \vdash \mathsf{match} \ p \ \mathsf{with} \mid x_1, \ldots, x_n \to e : \tau} \ (\mathsf{matchTuple}) \qquad \frac{\Gamma \vdash e_1 : \mathsf{int} \qquad \Gamma \vdash e_2 : \mathsf{int}}{\Gamma \vdash e_1 + e_2 : \mathsf{int}} \ (\mathsf{addInt})$$

$$\frac{\Gamma, x : \tau_1 \vdash e : \tau_2}{\Gamma \vdash \text{fun } x \to e : \tau_1 \to \tau_2} \text{ (fun)}$$

$$\frac{(v : \tau) \in \Gamma}{\Gamma \vdash v : \tau} \text{ (var)}$$



### Solution

#### Homework

Implement the function

```
val fill_in_steps: (int, int, int) -> int list
```

so that fill\_in\_steps (a, b, c) is the shortest list of numbers starting at a, ending at c, containing b and having consecutive adjacent elements

```
example: fill_in_steps (1, 4, -2) = [1;2;3;4;3;2;1;0;-1;-2]
```



#### Records

```
type point = { x_cord : float ; y_cord : float }
let origin : point = { x_cord = 0. ; y_cord = 0. }

type user = {
  name : string ;
  email : string ;
  num_posts : int ;
}
```

Records are *unordered labeled* fixed—length heterogeneous collections of data

They are useful for organizing large collections of data (akin to database records)

## Record Syntax (Informal)

For a record, we have to specify the type of each field When we construct a record, every field must have a value

#### Accessors

```
type point = { x_cord : float ; y_cord : float }
let dist (p : point) (q : point) =
  let xd = p.x_cord -. q.x_cord in
  let yd = p.y_cord -. q.y_cord in
  sqrt (xd *. xd +. yd *. yd)
```

Records support dot-notation

(we can also access records by pattern matching)

```
let new_post u : user =
    { u with num_posts = u.num_posts + 1 }
```

```
let new_post u : user =
    { u with num_posts = u.num_posts + 1 }
```

We can use with-syntax to update a smaller number of fields in a large record

```
let new_post u : user =
    { u with num_posts = u.num_posts + 1 }
```

We can use with-syntax to update a smaller number of fields in a large record

"u with num\_posts incremented, keep everything else the same"

```
let new_post u : user =
    { u with num_posts = u.num_posts + 1 }
```

We can use with-syntax to update a smaller number of fields in a large record

"u with num\_posts incremented, keep everything else the same"

**Data in functional languages are immutable.** This returns a new record with the update

# demo

(tuples and records)

#### Homework

- » Define formal syntax rules for records
- » Define formal typing rules for records
- >> Define formal semantics rules for records



# Unions

**Second Greatest Feature of OCaml** 

## Simple Variants

```
type os = BSD | Linux | MacOS | Windows
```

A **simple variant** is a *user-defined* type for values of a fixed collection of possibilities

First letter of type names is **lower\_case** and Constructor names is **Upper\_case** 

## Simple Variants

```
type os = BSD | Linux | MacOS | Windows
```

A **simple variant** is a *user-defined* type for values of a fixed collection of possibilities

First letter of type names is **lower\_case** and Constructor names is **Upper\_case** 

## Pattern Matching

```
let supported (sys : os) : bool =
  match sys with
    | BSD -> false
    | _ -> true
```

We work with variants by pattern matching:

» giving a <u>pattern</u> that a value can <u>match</u> with

>> writing what to do for each pattern

## Pattern Matching

We work with variants by pattern matching:

» giving a <u>pattern</u> that a value can <u>match</u> with

>> writing what to do for each pattern

## Data-Carrying Variants

Variants can carry data, which allows us to represent more complex structures

## Data-Carrying Variants

```
type linux distro = Arch | Fedora | NixOS | Ubuntu
           type os
             = BSD of int * int
             Linux of linux distro * int
            MacOS of int
Note the syntax | Windows of int
           let supported (sys : os) : bool =
             match sys with
             BSD (major, minor) -> major > 2 && minor > 3
             -> true
```

Variants can carry data, which allows us to represent more complex structures

#### Practice Problem

```
let area (s : shape) =
  match s with
  | Rect r -> r.base *. r.height
  | Triangle { sides = (a, b) ; angle } -> Float.sin angle *. a *. b
  | Circle r -> r *. r *. Float.pi
```

Define the variant **shape** which makes this function type-check



# What about variable length data?

# Recursive ADTs

**Greatest Feature of OCaml** 

## Example: Lists

```
type intlist
    = Nil
    | Cons of int * intlist

let example = Cons (1, Cons (2, Cons (3, Nil)))
```

The type **intlist** is available as the type of data which a constructor of **intlist** holds

We can use recursive ADTs to create variable—length data types

## Example: Lists

```
type intlist
    = Nil
    | Cons of int * intlist

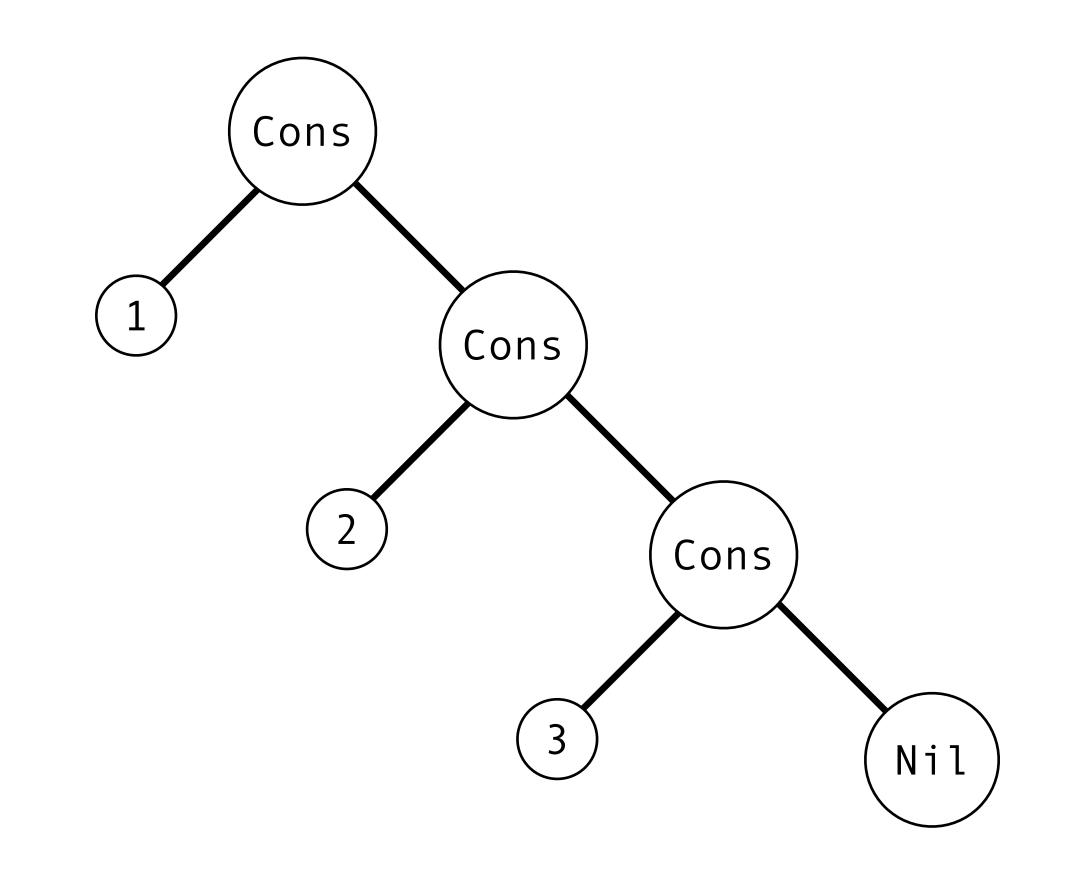
let example = Cons (1, Cons (2, Cons (3, Nil)))
```

The type **intlist** is available as the type of data which a constructor of **intlist** holds

We can use recursive ADTs to create variable—length data types

#### The Picture

```
Cons (1,
Cons (2,
Cons (3,
Nil))
```



We think of values of recursive variants as trees with constructors as nodes and carried data as leaves

# demo

(snoc for intlist)

### A More Interesting Example: Expressions

$$3 + ((2*4) - 14)$$

## A More Interesting Example: Expressions

$$3 + ((2*4) - 14)$$

Suppose we're building a calculator\*

## A More Interesting Example: Expressions

$$3 + ((2*4) - 14)$$

Suppose we're building a calculator\*

Before we compute the value of an input, we first have to find an abstract representation of the input

### A More Interesting Example: Expressions

$$3 + ((2*4) - 14)$$

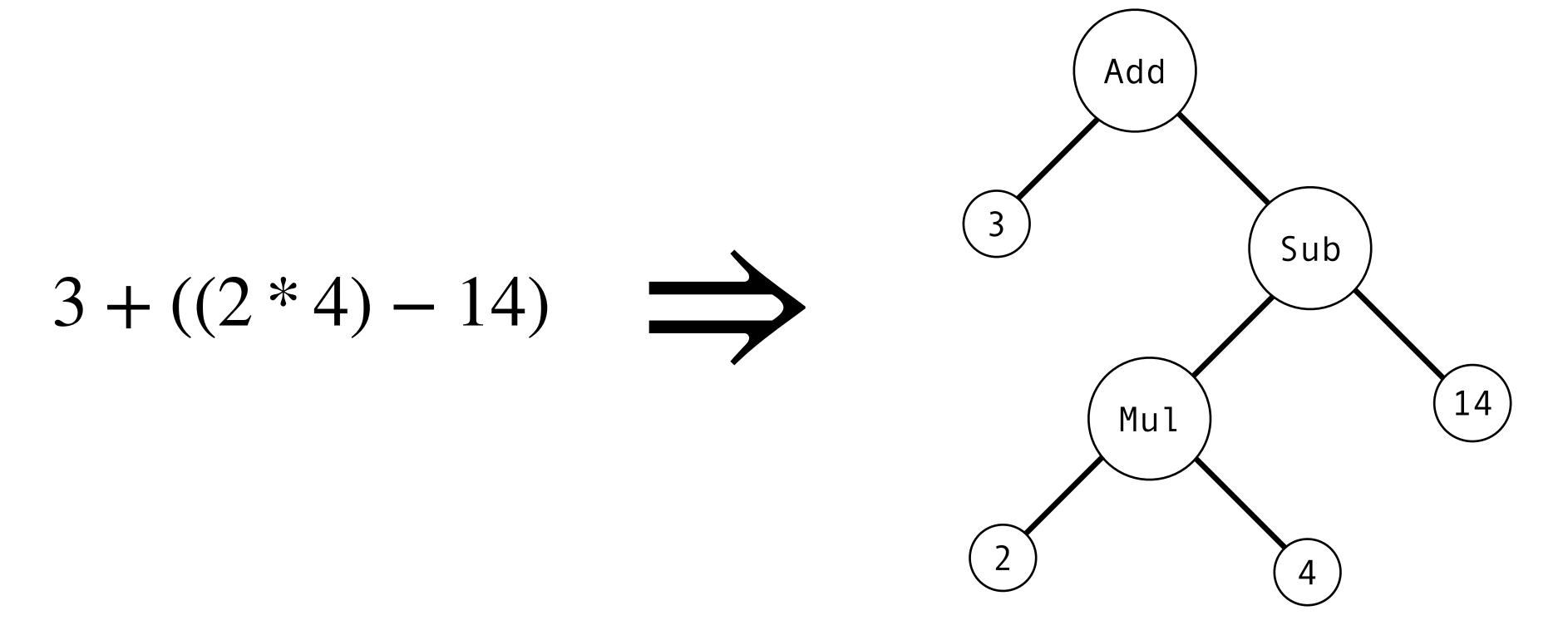
Suppose we're building a calculator\*

Before we compute the value of an input, we first have to find an abstract representation of the input

This will help us separate the tasks of evaluation and parsing

\*This is exactly what we'll be doing when we build an interpreter.

## A More Interesting Example: Expressions



We can represent an expression abstractly as a tree with operations as nodes and number values as leaves

#### A More Interesting Example: Expressions

```
type expr
    = Val of int
    | Add of expr * expr
    | Sub of expr * expr
    | Mul of expr * expr
    | Mul of expr * expr
let _ = Add (Val 3, Sub (Mul (Val 2, Val 4), Val 14))
```

Which means we can represent it as a recursive variant!

## Parametrized ADTs

The last piece of the puzzle: variants can be type agnostic

The last piece of the puzzle: variants can be type agnostic

This gives us a variant which is parametrically polymorphic

```
type 'a mylist
= Nil
    | Cons of 'a * 'a mylist

let e1 : int mylist = Cons (1, Cons (2, Cons (3, Nil)))
let e2 : string mylist = Cons ("1", Cons ("2", Cons ("3", Nil)))
```

The last piece of the puzzle: variants can be type agnostic

This gives us a variant which is parametrically polymorphic

```
type 'a mylist type constructor
= Nil
| Cons of 'a * 'a mylist

let e1 : int mylist = Cons (1, Cons (2, Cons (3, Nil)))
let e2 : string mylist = Cons ("1", Cons ("2", Cons ("3", Nil)))
```

The last piece of the puzzle: variants can be type agnostic

This gives us a variant which is parametrically polymorphic

## Parametric Polymorphism

```
let rev_tail (l : 'a list) : 'a list =
  let rec go acc l =
    match l with
    | [] -> acc
    | x :: xs -> go (x :: acc) xs
  in go [] l
```

## Parametric Polymorphism

```
let rev_tail (l : 'a list) : 'a list =
  let rec go acc l =
    match l with
    | [] -> acc
    | x :: xs -> go (x :: acc) xs
  in go [] l
```

This allows us to write functions which can be more generally applied (reversing a list does not depend on what's in the list)

## Parametric Polymorphism

```
let rev_tail (l : 'a list) : 'a list =
  let rec go acc l =
    match l with
    | [] -> acc
    | x :: xs -> go (x :: acc) xs
  in go [] l
```

This allows us to write functions which can be more generally applied (reversing a list does not depend on what's in the list)

Note. Because of type-inference, we rarely have to think about this

# We'll come back to this next time...

## Summary

Tuples, records, and ADTs help us organize data and create abstract interfaces

Recursive and parametrized ADTs give us richer structure