Lecture 21 – Algebraic Data Types (1)

COSE212: Programming Languages

Jihyeok Park



2023 Fall





- TFAE FAE with type system.
 - Type Checker and Typing Rules
 - Interpreter and Natural Semantics
- TRFAE RFAE with type system.
 - Type Checker and Typing Rules
 - Interpreter and Natural Semantics





- TFAE FAE with type system.
 - Type Checker and Typing Rules
 - Interpreter and Natural Semantics
- TRFAE RFAE with type system.
 - Type Checker and Typing Rules
 - Interpreter and Natural Semantics
- Let's learn algebraic data types (ADTs) and pattern matching!





- TFAE FAE with type system.
 - Type Checker and Typing Rules
 - Interpreter and Natural Semantics
- TRFAE RFAE with type system.
 - Type Checker and Typing Rules
 - Interpreter and Natural Semantics
- Let's learn algebraic data types (ADTs) and pattern matching!
- TAFAE TRFAE with ADTs and pattern matching.
 - Interpreter and Natural Semantics
 - Type Checker and Typing Rules

Recall



- TFAE FAE with type system.
 - Type Checker and Typing Rules
 - Interpreter and Natural Semantics
- TRFAE RFAE with type system.
 - Type Checker and Typing Rules
 - Interpreter and Natural Semantics
- Let's learn algebraic data types (ADTs) and pattern matching!
- TAFAE TRFAE with ADTs and pattern matching.
 - Interpreter and Natural Semantics
 - Type Checker and Typing Rules
- In this lecture, we will focus on Interpreter and Natural Semantics.

Contents



1. Algebraic Data Types (ADTs) and Pattern Matching

Recall: Types
Product Types
Union Types
Sum Types

Algebraic Data Types (ADTs)

Pattern Matching

2. TAFAE – TRFAE with ADTs and Pattern Matching

Concrete Syntax Abstract Syntax

3. Interpreter and Natural Semantics for TAFAE

Algebraic Data Types Function Application Pattern Matching Examples

Contents



1. Algebraic Data Types (ADTs) and Pattern Matching

Recall: Types
Product Types
Union Types
Sum Types
Algebraic Data Types (ADTs)
Pattern Matching

2. TAFAE – TRFAE with ADTs and Pattern Matching Concrete Syntax

3. Interpreter and Natural Semantics for TAFAE Algebraic Data Types
Function Application
Pattern Matching
Examples





A type is a set of values.

For example, the Int, Boolean, and Int => Int types are defined as the following sets of values in Scala.

```
Int = \{n \in \mathbb{Z} \mid -2^{31} \le n < 2^{31}\}
Boolean = \{\text{true}, \text{false}\}
Int \Rightarrow Int = \{f \mid f \text{ is a function from Int to Int}\}
```



A type is a set of values.

For example, the Int, Boolean, and Int => Int types are defined as the following sets of values in Scala.

```
Int = \{n \in \mathbb{Z} \mid -2^{31} \le n < 2^{31}\}
Boolean = \{\text{true}, \text{false}\}
Int \Rightarrow Int = \{f \mid f \text{ is a function from Int to Int}\}
```

Is it possible to define a **new type** by **combining** existing types?



A **type** is a set of values.

For example, the Int, Boolean, and Int => Int types are defined as the following sets of values in Scala.

```
Int = \{n \in \mathbb{Z} \mid -2^{31} \le n < 2^{31}\}
Boolean = \{\text{true}, \text{false}\}
Int => \text{Int} = \{f \mid f \text{ is a function from Int to Int}\}
```

Is it possible to define a **new type** by **combining** existing types? **Yes!**



A type is a set of values.

For example, the Int, Boolean, and Int => Int types are defined as the following sets of values in Scala.

```
\begin{array}{ll} \text{Int} &= \{n \in \mathbb{Z} \mid -2^{31} \leq n < 2^{31}\} \\ \text{Boolean} &= \{\texttt{true}, \texttt{false}\} \\ \text{Int => Int} &= \{f \mid f \text{ is a function from Int to Int}\} \end{array}
```

Is it possible to define a **new type** by **combining** existing types? **Yes!**

Product Types, Union Types, Sum Types, and Algebraic Data Types!

Product Types



Definition (Product Types)

A **product type** (τ_1, \ldots, τ_n) is a set of values of the form (v_1, \ldots, v_n) where τ_i is the type of v_i for $1 \le i \le n$.

Product Types



Definition (Product Types)

A **product type** (τ_1, \ldots, τ_n) is a set of values of the form (v_1, \ldots, v_n) where τ_i is the type of v_i for $1 \le i \le n$.

It is corresponds to the Cartesian product of sets:

$$(\tau_1,\ldots,\tau_n)=\tau_1\times\ldots\times\tau_n$$

Product Types



Definition (Product Types)

A **product type** (τ_1, \ldots, τ_n) is a set of values of the form (v_1, \ldots, v_n) where τ_i is the type of v_i for $1 \le i \le n$.

It is corresponds to the **Cartesian product** of sets:

$$(\tau_1,\ldots,\tau_n)=\tau_1\times\ldots\times\tau_n$$

For example, we can define product types in Scala as follows:



Definition (Union Types)

A **union type** $\tau_1 \mid \ldots \mid \tau_n$ is a set of values whose type is one of τ_1, \ldots, τ_n .



Definition (Union Types)

A **union type** $\tau_1 \mid \ldots \mid \tau_n$ is a set of values whose type is one of τ_1, \ldots, τ_n .

It is corresponds to the **union** of sets:

$$\tau_1 \mid \ldots \mid \tau_n = \tau_1 \cup \ldots \cup \tau_n$$



Definition (Union Types)

A union type $\tau_1 \mid \ldots \mid \tau_n$ is a set of values whose type is one of τ_1, \ldots, τ_n .

It is corresponds to the union of sets:

$$\tau_1 \mid \ldots \mid \tau_n = \tau_1 \cup \ldots \cup \tau_n$$

For example, we can define union types in Scala as follows:



Definition (Union Types)

A **union type** $\tau_1 \mid \ldots \mid \tau_n$ is a set of values whose type is one of τ_1, \ldots, τ_n .

It is corresponds to the union of sets:

$$\tau_1 \mid \ldots \mid \tau_n = \tau_1 \cup \ldots \cup \tau_n$$

For example, we can define union types in Scala as follows:

How can we **discriminate** between a square and a triangle?



Definition (Union Types)

A union type $\tau_1 \mid \ldots \mid \tau_n$ is a set of values whose type is one of τ_1, \ldots, τ_n .

It is corresponds to the union of sets:

$$\tau_1 \mid \ldots \mid \tau_n = \tau_1 \cup \ldots \cup \tau_n$$

For example, we can define union types in Scala as follows:

How can we discriminate between a square and a triangle? Sum types!

Sum Types



Definition (Sum Types)

A sum type $x_1(\tau_1) + \ldots + x_n(\tau_n)$ consists of variants $x_i(\tau_i)$ for $1 \le i \le n$. For each variant $x_i(\tau_i)$, x_i is the **constructor**, a function that takes a value v of type τ_i and generates a value $x_i(v)$ of the sum type.

Sum Types



Definition (Sum Types)

A sum type $x_1(\tau_1) + \ldots + x_n(\tau_n)$ consists of variants $x_i(\tau_i)$ for $1 \le i \le n$. For each variant $x_i(\tau_i)$, x_i is the **constructor**, a function that takes a value v of type τ_i and generates a value $x_i(v)$ of the sum type.

It is corresponds to a tagged union of sets:

$$x_1(\tau_1) + \ldots + x_n(\tau_n) = \{x_i(v) \mid \exists 1 \le i \le n. \text{ s.t. } v \in \tau_i\}$$

Sum Types



Definition (Sum Types)

A sum type $x_1(\tau_1) + \ldots + x_n(\tau_n)$ consists of variants $x_i(\tau_i)$ for $1 \le i \le n$. For each variant $x_i(\tau_i)$, x_i is the **constructor**, a function that takes a value v of type τ_i and generates a value $x_i(v)$ of the sum type.

It is corresponds to a tagged union of sets:

$$x_1(\tau_1) + \ldots + x_n(\tau_n) = \{x_i(v) \mid \exists 1 \le i \le n. \text{ s.t. } v \in \tau_i\}$$

For example, we can define **sum types** in Scala as follows:

Now, we can discriminate between a square and a triangle!





Definition (Sum Types)

A sum type $x_1(\tau_1) + \ldots + x_n(\tau_n)$ consists of variants $x_i(\tau_i)$ for $1 \le i \le n$. For each variant $x_i(\tau_i)$, x_i is the **constructor**, a function that takes a value v of type τ_i and generates a value $x_i(v)$ of the sum type.

Algebraic Data Types (ADTs)



Definition (Algebraic Data Types (ADTs))

An algebraic data type $x_1(\tau_{1,1},\ldots,\tau_{1,m_1})+\ldots+x_n(\tau_{n,1},\ldots,\tau_{n,m_n})$ is a recursive sum type of product types.

Algebraic Data Types (ADTs)



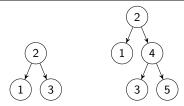
Definition (Algebraic Data Types (ADTs))

An algebraic data type $x_1(\tau_{1,1},\ldots,\tau_{1,m_1})+\ldots+x_n(\tau_{n,1},\ldots,\tau_{n,m_n})$ is a recursive sum type of product types.

For example, we can define algebraic data type for trees in Scala:

```
enum Tree:
   case Leaf(v: Int)
   case Node(1: Tree, v: Int, r: Tree)

val t1: Tree = Node(Leaf(1), 2, Leaf(3))
val t2: Tree = Node(Leaf(1), 2, Node(Leaf(3), 4, Leaf(5)))
```



Pattern Matching



Definition (Pattern matching)

We can use **pattern matching** for algebraic data types to identify which variant of the sum type a value belongs to and extract the data it contains.

Pattern Matching



Definition (Pattern matching)

We can use **pattern matching** for algebraic data types to identify which variant of the sum type a value belongs to and extract the data it contains.

For example, we can define a function sum that sums all the values in a tree using pattern matching (match) on the Tree type in Scala:

```
enum Tree:
    case Leaf(v: Int)
    case Node(1: Tree, v: Int, r: Tree)

def sum(t: Tree): Int = t match
    case Leaf(v) => v
    case Node(1, v, r) => sum(1) + v + sum(r)

sum(Node(Leaf(1), 2, Leaf(3))) // 6
sum(Node(Leaf(1), 2, Node(Leaf(3), 4, Leaf(5)))) // 15
```

Algebraic Data Types



Many functional languages support algebraic data types:

• Scala

```
enum Tree { Leaf(v: Int), Node(1: Tree, v: Int, r: Tree) }
```

• Haskell

```
data Tree = Leaf Int | Node Tree Int Tree
```

• Rust

```
enum Tree { Leaf(i32), Node(Tree, i32, Tree) }
```

OCaml

```
type tree = Leaf of int | Node of tree * int * tree
```

•

Contents



1. Algebraic Data Types (ADTs) and Pattern Matching

Recall: Types
Product Types
Union Types
Sum Types
Algebraic Data Types (ADTs)
Pattern Matching

2. TAFAE – TRFAE with ADTs and Pattern Matching

Concrete Syntax Abstract Syntax

3. Interpreter and Natural Semantics for TAFAE

Algebraic Data Types Function Application Pattern Matching Examples

TAFAE – TRFAE with ADTs and Pattern Matching PLRG

Now, let's extend TRFAE into TAFAE to support **algebraic data types** and **pattern matching**. (Assume that TRFAE supports multiple arguments for functions.)

```
/* TAFAE */
enum Tree {
  case Leaf(Number)
  case Node(Tree, Number, Tree)
}
Leaf(42) match {
  case Leaf(v) => v
  case Node(1, v, r) => v
}
```

For TAFAE, we need to extend expressions of TRFAE with

- 1 algebraic data types (ADTs)
- pattern matching
- 3 type names

Concrete Syntax



For TAFAE, we need to extend **expressions** of TRFAE with

- 1 algebraic data types (ADTs)
- pattern matching
- § type names

Concrete Syntax



For TAFAE, we need to extend expressions of TRFAE with

- 1 algebraic data types (ADTs)
- pattern matching
- **3** type names

We can extend the **concrete syntax** of FAE as follows:

```
// expressions
<expr> ::= ...
        | "enum" <id> "{" [ <variant> ";"? ]+ "}" ";"? <expr>
        | <expr> "match" "{" [ <mcase> ";"? ]+ "}"
// variants
<variant> ::= "case" <id> "(" ")"
           | "case" <id> "(" <type> [ ", " <type> ]* ")"
// match cases
<mcase> ::= "case" <id> "(" ")" "=>" <expr>
         | "case" <id> "(" <id> [ "," <id> ]* ")" "=>" <expr>
// types
```





```
enum Expr:
...
case TypeDef(name: String, varts: List[Variant], body: Expr)
case Match(expr: Expr, mcases: List[MatchCase])

case class Variant(name: String, ptys: List[Type]):
case class MatchCase(name: String, params: List[String], body: Expr):
enum Type:
...
case NameT(name: String)
```

Abstract Syntax



```
/* TAFAE */
enum Tree {
  case Leaf(Number)
  case Node(Tree, Number, Tree)
}
Leaf(42) match {
  case Leaf(v) => v
  case Node(1, v, r) => v
}
```

will be parsed to the following abstract syntax tree (AST) in Scala:

```
TypeDef("Tree", List(
    Variant("Leaf", List(NumT)),
    Variant("Node", List(NameT("Tree"), NumT, NameT("Tree")))
),
Match(App(Id("Leaf"), List(Num(42))), List(
    MatchCase("Leaf", List("v"), Id("v")),
    MatchCase("Node", List("l", "v", "r"), Id("v")))))
```

Contents



1. Algebraic Data Types (ADTs) and Pattern Matching

Recall: Types
Product Types
Union Types
Sum Types
Algebraic Data Types (ADTs
Pattern Matching

2. TAFAE - TRFAE with ADTs and Pattern Matching

Concrete Syntax Abstract Syntax

3. Interpreter and Natural Semantics for TAFAE

Algebraic Data Types Function Application Pattern Matching Examples

Interpreter and Natural Semantics for TAFAE



For TAFAE, we need to 1) implement the **interpreter** with environments:

```
def interp(expr: Expr, env: Env): Value = ???
```

and 2) define the **natural semantics** with environments:

$$\sigma \vdash e \Rightarrow v$$





For TAFAE, we need to 1) implement the **interpreter** with environments:

```
def interp(expr: Expr, env: Env): Value = ???
```

and 2) define the **natural semantics** with environments:

$$\sigma \vdash e \Rightarrow v$$

with a new kind of values called **constructor values** and **variant values**:

```
enum Value:
...
case ConstrV(name: String)
case VariantV(name: String, values: List[Value])
```

Algebraic Data Types



```
def interp(expr: Expr, env: Env): Value = expr match
   ...
   case TypeDef(_, ws, body) =>
    interp(body, env ++ ws.map(w => w.name -> ConstrV(w.name)))
```

$$\sigma \vdash e \Rightarrow v$$

$$\begin{array}{c} \text{TypeDef} \; \dfrac{\sigma[\mathsf{x}_1 \mapsto \langle \mathsf{x}_1 \rangle, \dots, \mathsf{x}_n \mapsto \langle \mathsf{x}_n \rangle] \vdash e \Rightarrow \mathsf{v} \\ \\ \sigma \vdash \mathsf{enum} \; t \; \left\{ \begin{array}{c} \mathsf{case} \; \mathsf{x}_1(\tau_{1,1}, \dots, \tau_{1,m_1}) \\ \dots \\ \mathsf{case} \; \mathsf{x}_n(\tau_{n,1}, \dots, \tau_{n,m_n}) \end{array} \right\}; \; e \Rightarrow \mathsf{v} \end{array}$$

```
/* TAFAE */
enum Tree { case Leaf(Number); case Node(Tree, Number, Tree) }
Leaf(42) match { case Leaf(v) => v; case Node(1, v, r) => v }
```

Algebraic Data Types



```
def interp(expr: Expr, env: Env): Value = expr match
   ...
   case App(f, es) => interp(f, env) match
      case CloV(ps, b, fenv) => ...
   case ConstrV(name) => VariantV(name, es.map(interp(_, env)))
   case v => error(s"not a function: ${v.str}")
```

$$\sigma \vdash e \Rightarrow v$$

$$\mathtt{App}_{\langle -\rangle} \; \frac{\sigma \vdash e_0 \Rightarrow \langle x \rangle \quad \sigma \vdash e_1 \Rightarrow v_1 \quad \dots \quad \sigma \vdash e_n \Rightarrow v_n}{\sigma \vdash e_0(e_1, \dots, e_n) \Rightarrow x(v_1, \dots, v_n)}$$

```
/* TAFAE */
enum Tree { case Leaf(Number); case Node(Tree, Number, Tree) }
Leaf(42) match { case Leaf(v) => v; case Node(1, v, r) => v }
```

Pattern Matching



```
def interp(expr: Expr, env: Env): Value = expr match
...
  case Match(expr, cases) => interp(expr, env) match
    case VariantV(wname, vs) => cases.find(_.name == wname) match
    case Some(MatchCase(_, ps, b)) =>
        if (ps.length != vs.length) error("arity mismatch")
        interp(b, env ++ (ps zip vs))
        case None => error(s"no such case: $wname")
        case v => error(s"not a variant: ${v.str}")
```

$$\sigma \vdash e \Rightarrow v$$

$$\begin{array}{c} 1 \leq i \leq n \quad \sigma \vdash e \Rightarrow x_i \big(v_1, \ldots, v_{m_i} \big) \quad \forall j < i. \; x_j \neq x_i \\ \frac{\sigma \big[x_{i,1} \mapsto v_1, \ldots, x_{i,m_i} \mapsto v_{m_i} \big] \vdash e_i \Rightarrow v \\ \hline \\ \sigma \vdash e \; \mathtt{match} \; \left\{ \begin{array}{c} \mathtt{case} \; x_1 \big(x_{1,1}, \ldots, x_{1,m_1} \big) \Rightarrow e_1 \\ \ldots \\ \mathtt{case} \; x_n \big(x_{n,1}, \ldots, x_{n,m_n} \big) \Rightarrow e_n \end{array} \right\} \Rightarrow v \end{array}$$





There exists an order between the match cases: first match wins!

$$\text{Match} \begin{array}{c} 1 \leq i \leq n & \sigma \vdash e \Rightarrow x_i(v_1, \ldots, v_{m_i}) & \forall j < i. \; x_j \neq x_i \\ \sigma[x_{i,1} \mapsto v_1, \ldots, x_{i,m_i} \mapsto v_{m_i}] \vdash e_i \Rightarrow v \\ \hline \sigma \vdash e \; \text{match} \; \left\{ \begin{array}{c} \text{case} \; x_1(x_{1,1}, \ldots, x_{1,m_1}) \Rightarrow e_1 \\ \ldots \\ \text{case} \; x_n(x_{n,1}, \ldots, x_{n,m_n}) \Rightarrow e_n \end{array} \right\} \Rightarrow v \end{array}$$

Example 1



```
/* TAFAE */
enum A { case B(Boolean); case C(Number) }
C(42) match { case B(b) => b; case C(n) => n < 0 }</pre>
```

Match

$$\sigma_1 \vdash C(42) \text{ match } \left\{ \begin{array}{l} \text{case } B(b) \Rightarrow b \\ \text{case } C(n) \Rightarrow n < 0 \end{array} \right\} \Rightarrow \text{false}$$

TypeDef

where

$$\begin{array}{rcl}
\sigma_1 & = & [B \mapsto \langle B \rangle, C \mapsto \langle C \rangle] \\
\sigma_2 & = & \sigma_2[n \mapsto 42]
\end{array}$$

Example 2



In **TFAE**, we cannot define mkRec because of the lack of **recursive types** in the language:

```
/* TFAE */
val mkRec = (body: (Number => Number) => Number => Number) => {
  val fX = (fY: ???) => {
    val f = (x: Number) \Rightarrow fY(fY)(x):
    body(f)
  };
  fX(fX)
};
val sum = mkRec((sum: Number => Number) => (n: Number) =>
  if (n < 1) 0
  else n + sum(n + -1):
sum(10)
```

Example 2



Now, we can define mkRec in **TAFAE** because **algebraic data types** are **recursive types**:

```
/* TAFAE */
enum T { case T(T => Number => Number) }
val mkRec = (body: (Number => Number) => Number => Number) => {
  val fX = (fY: T) \Rightarrow {
    val f = (x: Number) \Rightarrow fY match \{ case T(fZ) \Rightarrow fZ(fY)(x) \};
    body(f)
  }:
  fX(T(fX))
};
val sum = mkRec((sum: Number => Number) => (n: Number) =>
  if (n < 1) 0
  else n + sum(n + -1):
sum(10)
```









We can define list type as well using ADTs in TAFAE:





We can define list type as well using ADTs in TAFAE:

However, it only works for monomorphic lists (i.e., lists of numbers)





We can define list type as well using ADTs in TAFAE:

However, it only works for monomorphic lists (i.e., lists of numbers)

We will learn parametric polymorphism later in this course.

Summary



1. Algebraic Data Types (ADTs) and Pattern Matching

Recall: Types
Product Types

Union Types

Sum Types

Algebraic Data Types (ADTs)

Pattern Matching

2. TAFAE – TRFAE with ADTs and Pattern Matching

Concrete Syntax

Abstract Syntax

3. Interpreter and Natural Semantics for TAFAE

Algebraic Data Types

Function Application

Pattern Matching

Examples

Next Lecture



• Algebraic Data Types (2)

Jihyeok Park
jihyeok_park@korea.ac.kr
https://plrg.korea.ac.kr