

3.3 Type Inference

In the first section of this chapter we defined a type system, in the second section we introduced a syntax for our language. Now we want to define rules how to obtain the type of a given program written in our language.

3.3.1 Typing Judgments

A type system is a set of inference rules to derive various kinds of typing judgments. These *inference rules* have the following form

$$\frac{P_1 \dots P_n}{C}$$

where the judgments P_1 up to P_n are the premises of the rule and the judgment C is its conclusion.

We can read it in two ways:

- “If all premises hold then the conclusion holds as well” or
- “To prove the conclusion we need to prove all premises”.

We will now provide a judgment for every production rule defined in the last section. Ultimately, we will have a judgment $p : T$ which indicates that a program p is of a type T and therefore well-formed.

If the type T is known then we talk about *type checking* else we call the process of finding the judgment *type inference*.

TYPE SIGNATURE JUDGMENTS

For type signature judgments, let Γ be a type context, $T \in \mathcal{T}$ and $a_i \in \mathcal{V}, T_i \in \mathcal{T}$ for all $i \in \mathbb{N}_1^n$ and $n \in \mathbb{N}$.

For $ltf \in \langle \text{list-type-fields} \rangle$ the judgment has the form

$$\Gamma \vdash ltf : \{a_1 : T_1, \dots, a_n : T_n\}$$

which can be read as “given Γ , ltf has the type $\{a_1 : T_1, \dots, a_n : T_n\}$ ”.

For $lt \in \langle \text{list-type} \rangle$ the judgment has the form

$$\Gamma \vdash lt : (T_1, \dots, T_n)$$

which can be read as “given Γ , lt defines the list (T_1, \dots, T_n) ”.

For $t \in \langle \text{type} \rangle$ the judgment has the form

$$\Gamma \vdash t : T$$

which can be read as “given Γ , t has the type T ”.

EXPRESSION JUDGMENTS

For expression judgments, let Γ, Δ be type contexts, $T \in \mathcal{T}$, $a \in \mathcal{V}$ and $T_i \in \mathcal{T}, a_i \in \mathcal{V}$ for all $i \in \mathbb{N}_0^n, n \in \mathbb{N}$.

For $lef \in \langle \text{list-exp-field} \rangle$ the judgment has the form

$$\Gamma, \Delta \vdash lef : \{a_1 : T_1, \dots, a_n : T_n\}$$

which can be read as “given Γ and Δ , lef has the type $\{a_1 : T_1, \dots, a_n : T_n\}$ ”.

For $mes \in \langle \text{maybe-exp-sign} \rangle$ the judgment has the form

$$\Gamma, mes \vdash a : T$$

which can be read as “given Γ , a has the type T under the assumption mes ”.

For $b \in \langle \text{bool} \rangle$ the judgment has the form

$$b : T$$

which can be read as “ b has the type T ”.

For $i \in \langle \text{int} \rangle$ the judgment has the form

$$e : T$$

which can be read as “ i has the type T ”.

For $le \in \langle \text{list-exp} \rangle$ the judgment has the form

$$\Gamma, \Delta \vdash le : List\ T$$

which can be read as “given Γ and Δ , le has the type $List\ T$ ”.

For $e \in \langle \text{exp} \rangle$ the judgment has the form

$$\Gamma, \Delta \vdash e : T$$

which can be read as “given Γ and Δ , e is of type T ”.

STATEMENT JUDGMENTS

For statement judgments, let $\Gamma, \Gamma_1, \Gamma_2, \Delta, \Delta_1, \Delta_2$ be a type contexts, $T, T_1, T_2 \in \mathcal{T}$, $a \in \mathcal{V}$ and $T_i, A_i \in \mathcal{T}, a_i \in \mathcal{V}$ for $i \in \mathbb{N}_0^n$ and $T_{i,j} \in \mathcal{T}$ for $i \in \mathbb{N}_0^n, n \in \mathbb{N}, j \in \mathbb{N}_0^{k_i}$ and $k_i \in \mathbb{N}$.

For $lsv \in \langle \text{list-statement-var} \rangle$ the judgment has the form

$$lsv : (a_1, \dots, a_n)$$

which can be read as “ lsv describes the list (a_1, \dots, a_n) ”.

For $ls \in \langle \text{list-statement} \rangle$ the judgment has the form

$$\Gamma_1, \Delta_2, ls \vdash \Gamma_2, \Delta_2$$

which can be read as “the list of statements ls maps Γ_1 to Γ_2 and Δ_1 to Δ_2 ”.

For $mss \in \langle \text{maybe-statement-sign} \rangle$ the judgment has the form

$$\Gamma, mss \vdash a : T$$

which can be read as “given Γ , a has the type T_2 under the assumption mss ”.

For $s \in \langle \text{statement} \rangle$ the judgment has the form

$$\Gamma_1, \Delta_1, s \vdash \Gamma_2, \Delta_2$$

which can be read as “the statement s maps Γ_1 to Γ_2 and Δ_1 to Δ_2 ”.

For $mms \in \langle \text{maybe-main-sign} \rangle$ the judgment has the form

$$\Gamma, mms \vdash \text{main} : T$$

which can be read as “the main function has type T under the assumption mms ”.

For $prog \in \langle \text{program} \rangle$ the judgment has the form

$$prog : T$$

which can be read as “the program $prog$ is wellformed and has the type T ”.

3.3.2 Auxiliary Definitions

We will assume that T is a mono type and T is a type variable is defined. $T_1 = T_2$ denotes the equality of two given types T_1 and T_2 .

We will write $a_1, \dots, a_n = \text{free}(T)$ to denote all free variables a_1, \dots, a_n of T .

INSTANTIATION, GENERALIZATION

The type system that we are using is polymorphic, meaning that whenever a judgment holds for a type, it will also hold for a more specific type. To counter this polymorphism we will force the types in a judgment to be unique by explicitly stating whenever we want to use a more specific or general type.

Definition 3.1: Most General Type

Let Γ be a type context, $T \in \mathcal{T}$, .

We define $\bar{\Gamma} : \Gamma \rightarrow \mathcal{T}$ as

$$\begin{aligned} \bar{\Gamma}(T) &:= \forall a_1 \dots \forall a_n. T_0 \\ &\text{such that } \{a_1, \dots, a_n\} = \text{free}(T') \setminus \{a \mid (a, _) \in \Gamma\} \\ &\text{where } a_i \in \mathcal{V} \text{ for } i \in \mathbb{N}_0^n \text{ and } T_0 \text{ is the mono type of } T. \end{aligned}$$

We say $\bar{\Gamma}(T)$ is *the most general type* of T .

The most general type ensures that all type variables are bound by either a quantifier or a type alias in the type context Γ . It also ensure that every type variable bound by a quantifier occurs in the mono type T_0 .

The act of replacing types with more general ones, by binding free variables, is called *Generalization* and the act of replacing are more general type with a more specific type is called *Instantiation*. Both rules are typically in the text books introduced as an additional inference rule.

PREDEFINED TYPES

Additionally, we define

$$\begin{aligned} Bool &:= \mu_ . True | False \\ Nat &:= \mu C . 1 | Succ\ C \\ Int &:= \mu_ . 0 \mid Pos\ Nat \mid Neg\ Nat \\ List &:= \forall a . \mu C . [\] \mid Cons\ a\ C \end{aligned}$$

3.3.3 Inference Rules for Type Signatures

LIST-TYPE-FIELDS

Judgment: $\Gamma \vdash ltf : \{a_1 : T_1, \dots, a_n : T_n\}$

$$\Gamma \vdash "" : \{\}$$

$$\frac{\Gamma \vdash t : T_0 \quad \Gamma \vdash ltf : \{a_1 : T_1, \dots, a_n : T_n\} \quad \{a_0 : T_0, a_1 : T_1, \dots, a_n : T_n\} = T}{\Gamma \vdash a_0 \text{ " : " } t \text{ " , " } ltf : T}$$

The type context Γ is used in the judgment $\Gamma \vdash t : T_0$ that turns the type signature t into a type T_0 .

LIST-TYPE

Judgment: $\Gamma \vdash lt : (T_1, \dots, T_n)$

$$\Gamma \vdash "" : ()$$

$$\frac{\Gamma \vdash t : T_0 \quad \Gamma \vdash lt : (T_1, \dots, T_n) \quad (T_0, T_1, \dots, T_n) = T}{\Gamma \vdash t \text{ lt : } T}$$

TYPE

Judgment: $\Gamma \vdash t : T$

$$\frac{Bool = T}{\Gamma \vdash \text{"Bool"} : T}$$

$$\frac{Int = T}{\Gamma \vdash \text{"Int"} : T}$$

$$\frac{List\ T_2 = T_1 \quad \Gamma \vdash t : T_2}{\Gamma \vdash \text{"List"}\ t : T_1}$$

$$\frac{(T_1, T_2) = T_0 \quad \Gamma \vdash t_1 : T_1 \quad \Gamma \vdash t_2 : T_2}{\Gamma \vdash \text{"("}\ t_1\ \text{"},\ \text{"}\ t_2\ \text{"})" : T_0}$$

$$\frac{\Gamma \vdash ltf : T}{\Gamma \vdash \text{"{"}\ ltf\ \text{"}"} : T}$$

$$\frac{T_1 \rightarrow T_2 = T_0 \quad \Gamma \vdash t_1 : T_1 \quad \Gamma \vdash t_2 : T_2}{\Gamma \vdash t_1 \rightarrow t_2 : T_0}$$

$$\frac{(c, T') \in \Gamma \quad \Gamma \vdash l : (T_1, \dots, T_n) \quad \overline{T'}\ T_1 \dots T_n = T}{\Gamma \vdash c\ l : T}$$

For a given type T we write the application constructor as \overline{T} .

$$\frac{\forall a.a = T}{\Gamma \vdash a : T}$$

Example 3.1

In example ?? we have looked at the syntax for a list reversing function.

The type signature for the `reverse` function was `List a -> List a`. We will now show how we can obtain the corresponding type T_0 . For that, let $\Gamma = \emptyset$.

$$\frac{\frac{\frac{\overline{\forall a.a = T_3}}{\emptyset \vdash a : T_3} \quad \overline{List\ T_3 = T_1}}{\emptyset \vdash Lista : T_1} \quad \frac{\frac{\frac{\overline{\forall a.a = T_4}}{\emptyset \vdash a : T_4} \quad \overline{List\ T_4 = T_2}}{\emptyset \vdash Lista : T_2}}{T_1 \rightarrow T_2 = T_0}}{\emptyset \vdash \text{List}\ a \rightarrow \text{List}\ a : T_0}$$

We can therefore conclude that $T_0 = List\ (\forall a.a) \rightarrow List\ (\forall a.a) = \forall a.List\ a \rightarrow List\ a$.

3.3.4 Inference Rules for Expressions

LIST-EXP-FIELD

Judgment: $\Gamma, \Delta \vdash lef : \{a_1 : T_1, \dots, a_n : T_n\}$

$$\frac{\Gamma, \Delta \vdash e : T}{\Gamma, \Delta \vdash a "=" e : \{a : T\}}$$

$$\frac{\Gamma, \Delta \vdash lef : T \quad \Gamma, \Delta \vdash e : T_0 \quad \{a_0 : T_0, \dots, a_n : T_n\} = T}{\Gamma, \Delta \vdash a_0 "=" e ", " lef : T}$$

MAYBE-EXP-SIGN

Judgment: $\Gamma, mes \vdash a : T$

$$\Gamma, "" \vdash a : T$$

If no argument is given, then we do nothing.

$$\frac{\Gamma \vdash t : T a_1 = a_2}{\Gamma, a_1 ":" "t"; " \vdash a_2 : T}$$

If we have a variable a_1 and a type T , then the variables a_2 need to match. The type signature t defines the type of a_2 .

BOOL

Judgment: $b : T$

$$b : Bool$$

INT

Judgment: $i : T$

$$i : Int$$

We have proven in theorem ?? that *Nat* is isomorph to \mathbb{N} . Is should be trivial to therefore conclude that *Int* is isomorph to \mathbb{Z} . And therefore this rule is justified.

LIST-EXP

Judgment: $\Gamma, \Delta \vdash le : List T$

$$\Gamma, \Delta \vdash "" : \forall a. List\ a$$

$$\frac{\Gamma, \Delta \vdash e : T \quad \Gamma, \Delta \vdash le : List\ T}{\Gamma, \Delta \vdash e\ ",\ " le : List\ T}$$

EXP

Judgment: $\Gamma, \Delta \vdash e : T$

$$\Gamma, \Delta \vdash "foldl" : \forall a. \forall b. (a \rightarrow b \rightarrow b) \rightarrow b \rightarrow List\ a \rightarrow b$$

$$\Gamma, \Delta \vdash "(::)" : \forall a. a \rightarrow List\ a \rightarrow List\ a$$

$$\Gamma, \Delta \vdash "(+)" : Int \rightarrow Int \rightarrow Int$$

$$\Gamma, \Delta \vdash "(-)" : Int \rightarrow Int \rightarrow Int$$

$$\Gamma, \Delta \vdash "(*)" : Int \rightarrow Int \rightarrow Int$$

$$\Gamma, \Delta \vdash "(//)" : Int \rightarrow Int \rightarrow Int$$

$$\Gamma, \Delta \vdash "<" : Int \rightarrow Int \rightarrow Bool$$

$$\Gamma, \Delta \vdash "(==)" : Int \rightarrow Int \rightarrow Bool$$

$$\Gamma, \Delta \vdash "not" : Bool \rightarrow Bool$$

$$\Gamma, \Delta \vdash "(&\&)" : Bool \rightarrow Bool \rightarrow Bool$$

$$\Gamma, \Delta \vdash "(||)" : Bool \rightarrow Bool \rightarrow Bool$$

$$\frac{\Gamma, \Delta \vdash e_1 : T_1 \quad \Gamma, \Delta \vdash e_2 : T_1 \rightarrow T_2}{\Gamma, \Delta \vdash e_1\ "|\>"\ e_2 : T_2}$$

$$\frac{\Gamma, \Delta \vdash e_1 : T_1 \rightarrow T_2 \quad \Gamma, \Delta \vdash e_2 : T_2 \rightarrow T_3}{\Gamma, \Delta \vdash e_1\ ">>"\ e_2 : T_1 \rightarrow T_3}$$

$$\frac{\Gamma, \Delta \vdash e_1 : Bool \quad \Gamma, \Delta \vdash e_2 : T \quad \Gamma, \Delta \vdash e_3 : T}{\Gamma, \Delta \vdash \text{"if" } e_1 \text{ "then" } e_2 \text{ "else" } e_3 : T}$$

$$\frac{\Gamma, \Delta \vdash lef : \{a_1 : T_1, \dots, a_n : T_n\}}{\Gamma, \Delta \vdash \text{"{" } lef \text{ "}" } : \{a_1 : T_1, \dots, a_n : T_n\}}$$

$$\Gamma, \Delta \vdash \text{"{"} : \{\}$$

$$\frac{\Gamma, \Delta \vdash lef : \{a_1 : T_1, \dots, a_n : T_n\} \quad \Gamma, \Delta \vdash (a, \bar{\Gamma}(T_0)) \in \Delta \quad T_0 = \{a_1 : T_1, \dots, a_n : T_n, \dots\}}{\Gamma, \Delta \vdash \text{"{" } a \text{ " | " } lef \text{ "}" } : T_0}$$

Since Elm version 0.19, released in 2018, setters are not allowed to change the type of a field in a record.

$$\frac{(a_1, \{a_2 : T, \dots\}) \in \Delta}{\Gamma, \Delta \vdash a_1 \text{"."} a_2 : T}$$

$$\frac{(a, _) \notin \Delta \quad \Gamma, \Delta \vdash e_1 : T_1 \quad mes : T_1 \vdash a : T_1 \quad \Gamma, \Delta \cup \{(a, \bar{\Gamma}(T_1))\} \vdash e_2 : T_2}{\Gamma, \Delta \vdash \text{"let" } mes \text{ a "=" } e_1 \text{ "in" } e_2 : T_2}$$

$$\frac{\Gamma, \Delta \vdash e_1 : T_1 \rightarrow T_2 \quad \Gamma, \Delta \vdash e_2 : T_1}{\Gamma, \Delta \vdash e_1 \ e_2 : T_2}$$

$$\frac{b : T}{\Gamma, \Delta \vdash b : T}$$

$$\frac{i : T}{\Gamma, \Delta \vdash i : T}$$

$$\frac{\Gamma, \Delta \vdash le : T}{\Gamma, \Delta \vdash \text{"[" } le \text{ "]" } : T}$$

$$\frac{\Gamma, \Delta \vdash e_1 : T_1 \quad \Gamma, \Delta \vdash e_2 : T_2}{\Gamma, \Delta \vdash \text{"(" } e_1 \text{ ", " } e_2 \text{ ")" } : (T_1, T_2)}$$

$$\frac{\Gamma, \Delta \cup \{(a, \bar{\Gamma}(T_1))\} \vdash e : T_2}{\Gamma, \Delta \vdash \text{"\ " } a \text{ "->" } e : T_1 \rightarrow T_2}$$

$$\frac{(c, \bar{\Gamma}(T)) \in \Delta}{\Gamma, \Delta \vdash c : T}$$

$$\frac{(a, \bar{\Gamma}(T)) \in \Delta}{\Gamma, \Delta \vdash a : T}$$

Example 3.2

In example ?? we have looked at the syntax for a list reversing function. We can now check the type $T_0 = \forall a. \text{List } a \rightarrow \text{List } a$ of the **reverse** function for $\Gamma = \Delta = \emptyset, \Delta = \emptyset$. The body of the *reverse* function is as follows:

`foldl (::) []`

$$\frac{\frac{\frac{\overline{\emptyset, \emptyset \vdash \text{"foldl"} : T_2} \quad \overline{\emptyset, \emptyset \vdash \text{" (::)"} : \forall a. \text{List } a \rightarrow \text{List } a}}{\emptyset, \emptyset \vdash \text{"foldl (::)"} : T_1} \quad \frac{\overline{\emptyset, \emptyset \vdash \text{" " } : \forall a. a}}{\emptyset, \emptyset \vdash \text{" [] " } : \forall a. \text{List } a}}{\emptyset, \emptyset \vdash \text{"foldl (::) []" } : T_0}$$

where $T_1 = \forall a. \text{List } a \rightarrow \text{List } a \rightarrow \text{List } a$ and $T_2 = \forall a. (\text{List } a \rightarrow \text{List } a) \rightarrow \text{List } a \rightarrow \text{List } a \rightarrow \text{List } a$.

3.3.5 Inference Rules for Statements

LIST-STATEMENT-VAR

Judgment: $lsv : (a_1, \dots, a_n)$

`" " : ()`

$$\frac{lsv : (a_1, \dots, a_n)}{a_0 \text{ } lsv : (a_0, a_1, \dots, a_n)}$$

LIST-STATEMENT-SORT

Judgment: $lss : (c_1 : (T_{1,1}, \dots, T_{1,k_1}), \dots, c_n : (T_{n,1}, \dots, T_{n,k_n}))$

$$\frac{\Gamma \vdash lt : (T_0, \dots, T_n)}{c \text{ } lt : (c : (T_0, \dots, T_n))}$$

$$\frac{\Gamma \vdash lt : (T_{0,1}, \dots, T_{0,k_n}) \quad lss : \begin{pmatrix} a_1 : (T_{1,1}, \dots, T_{1,k_1}), \\ \vdots \\ a_n : (T_{n,1}, \dots, T_{n,k_n}) \end{pmatrix}}{c \text{ } lt \text{ " | " } lss : \begin{pmatrix} a_0 : (T_{0,1}, \dots, T_{0,k_0}), \\ a_1 : (T_{1,1}, \dots, T_{1,k_1}), \\ \vdots \\ a_n : (T_{n,1}, \dots, T_{n,k_n}) \end{pmatrix}}$$

LIST-STATEMENT

Judgment: $\Gamma_1, \Delta_1, ls \vdash \Gamma_2, \Delta_2$

$$\frac{\Gamma_1 = \Gamma_2 \quad \Delta_1 = \Delta_2}{\Gamma_1, \Delta_1 \text{""} \vdash \Gamma_2, \Delta_2}$$

$$\frac{\Gamma_1, \Delta_1, s \vdash \Gamma_2, \Delta_2 \quad \Gamma_2, \Delta_2, ls \vdash \Gamma_3, \Delta_3}{\Gamma_1, \Delta_1, s \text{"}; " ls \vdash \Gamma_3, \Delta_3}$$

MAYBE-STATEMENT-SIGN

Judgment: $\Gamma, mss \vdash a : T$

$$\Gamma, \text{""} \vdash a : T$$

$$\frac{\Gamma \vdash t : T a_1 = a_2}{\Gamma, a_1 \text{" : " } t \text{" ; " } \vdash a_2 : T}$$

STATEMENT

Judgment: $\Gamma_1, \Delta_1, s \vdash \Gamma_2, \Delta_2$

$$\frac{\Gamma_1 = \Gamma_2 \quad (a, _) \notin \Delta_1 \quad \Gamma_1, mss \vdash e : T \quad \Gamma_1, \Delta_1 \vdash e : T \quad \Delta_2 = \Delta_1 \cup \{(a, \bar{\Gamma}(T))\}}{\Gamma_1, \Delta_1, mss \ a \text{"="} e \vdash \Gamma_2, \Delta_2}$$

$$\frac{\Delta_1 = \Delta_2 \quad (c, _) \notin \Gamma_1 \quad \Gamma \vdash t : T_1 \quad T_2 \text{ is a mono type} \quad lsv : (a_1, \dots, a_n) \quad \{a_1 \dots a_n\} = \text{free}(T_2) \quad \forall a_1 \dots \forall a_n. T_2 = T_1 \quad \Gamma_2 = \Gamma_1 \cup \{(c, T_1)\}}{\Gamma_1, \Delta_1, \text{"type alias"} \ c \ lsv \text{"="} t \vdash \Gamma_2, \Delta_2}$$

$$\frac{\begin{array}{l} (c, _) \notin \Gamma_1 \quad lsv : (a_1, \dots, a_n) \\ lss : (c_1 : (T_{1,1}, \dots, T_{1,k_1}), \dots, c_n : (T_{n,1}, \dots, T_{n,k_n})) \\ \Delta_1 \cap \{(c_1, _), \dots, (c_n, _)\} = \emptyset \quad \{a_1 \dots a_n\} = \text{free}(T_2) \\ \mu C. c_1 \ T_{1,1} \ \dots \ T_{1,k_1} \mid \dots \mid c_n \ T_{n,1} \ \dots \ T_{n,k_n} = T_2 \quad \forall a_1 \dots \forall a_n. T_2 = T_1 \\ \Gamma_1 \cup \{(c, T_1)\} = \Gamma_2 \quad \Delta_1 \cup \left\{ \begin{array}{l} (c_1, \bar{\Gamma}(T_{1,1} \rightarrow \dots \rightarrow T_{1,k_1} \rightarrow T_1)), \\ \vdots \\ (c_n, \bar{\Gamma}(T_{n,1} \rightarrow \dots \rightarrow T_{n,k_n} \rightarrow T_1)) \end{array} \right\} = \Delta_2 \end{array}}{\Gamma_1, \Delta_1, \text{"type"} \ c \ lsv \text{"="} lss \vdash \Gamma_2, \Delta_2}$$

The list lss provides us with the structure of the type. From there we construct the type T_2 and bind all variables, thus creating the poly type T_1 . Additionally, every

sort c_i for $i \in \mathbb{N}_1^n$ has its own constructor that gets added to Δ_1 under the name c_i . In Elm these constructors are the only constants beginning with an upper-case letter.

MAYBE-MAIN-SIGN

Judgment: $\Gamma, mms \vdash \text{main} : T$

$$\Gamma, "" \vdash \text{main} : T$$

$$\frac{\Gamma \vdash t : T}{\Gamma, \text{"main" : "t";} \vdash \text{main} : T}$$

PROGRAM

Judgment: $prog : T$

$$\frac{\emptyset, \emptyset, ls \vdash \Gamma, \Delta \quad \Gamma, mms \vdash \text{main} : T \quad \Gamma, \Delta \vdash e : T}{ls \ mms \ \text{"main" = " e : T}}$$