Type checker for System F

Mudathir Mahgoub

May 3, 2018

1 Project Problem

This project is a type checker for annotated simply typed lambda calculus and system F. It supports subtyping for simple types. In general, a type checker would decide whether $\Gamma \vdash t : T$ is derivable: can the term t be assigned the type T under the typing context Γ ?

Section 2 describes the software and its architecture. Section 3 describes the rules used for simple types and provides some examples. It also explains how subsumption rule can be delayed until variable terms are generated, and provides a proof for correctness. Section 3 describes the rules used for system F and provides some examples.

2 Software description

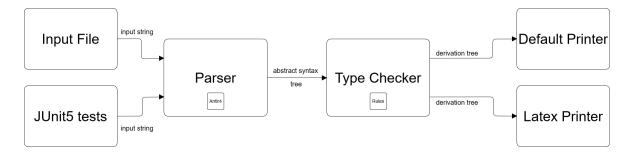


Figure 1: Project architecture.

The project is implemented using Java and the executable is a jar file (target/TypeChecker.jar) which is generated using the command:

mvn install

The program receives as an input a text file containing subtypes definitions and a judgment to be checked. For testing, **JUnit5** was used to test the program directly without files. The input is passed to the **Parser** which uses the ANTLR library to parse the input into an abstract syntax tree. This abstract syntax tree is consumed by the **Type Checker** which uses the rules in sections 3 and 4 to build a derivation tree and determine the answer of the type checking. The answer can be **Yes**, **No**, or **Unknown**. Finally the derivation tree can be printed using the **Default Printer** or the **Latex Printer**.

Here is an example of an input:

```
Listing 1: test.txt
```

```
SubBase(bool, int); \\ . \mid - \lambda x. \mid x. \mid (x y)[bool]: (int ->T) -> (bool -> T);
```

Here is the output of the default printer:

Listing 2: java -jar TypeChecker.jar -i test.txt

Here is the output of the latex printer which uses bussproofs package centered proofs:

Listing 3: java -jar TypeChecker.jar -i test.txt -latex

```
Yes
\begin{prooftree}
\AxiomC{} \RightLabel{\scriptsize var}
\UnaryInfC{$x: bool \vdash x : bool$}
\AxiomC{\scriptsize $SubBase(bool, int)$}
\RightLabel{\scriptsize subBase}
\UnaryInfC{$bool <: int$}
\RightLabel{\scriptsize subsumption}
\BinaryInfC{$x: bool \vdash x : int$}
\end{prooftree}</pre>
```

The following sections focus on terms and rules used in the **Type Checker**.

3 Simple types

Annotated simply typed terms recognized by the parser have the form:

$$t ::= x \mid (t_1 t_2)[T] \mid \lambda x.t$$

The annotation [T] in the application term $(t_1t_2)[T]$ is used to remove the non-determinism in the application rule. This annotation is required in the parser.

3.1 Typing rules for simple types

1.
$$\frac{\Gamma(x) = T}{\Gamma \vdash x : T} \text{ var}$$
2.
$$\frac{\Gamma \vdash t_1 : T_1 \to T_2 \qquad \Gamma \vdash t_2 : T_1}{\Gamma \vdash (t_1 t_2)[T_1] : T_2} \text{ app}$$
3.
$$\frac{\Gamma, x : T_1 \vdash t : T_2}{\Gamma \vdash \lambda x . t : T_1 \to T_2} \lambda$$

3.2 Subtyping rules for simple types

1.
$$\frac{\Gamma \vdash t : T_1 \qquad T_1 <: T_2}{\Gamma \vdash t : T_2}$$
 subsumption
$$2 \cdot \frac{T <: T}{T <: T}$$
 reflexive

3.
$$\frac{SubBase(b_1, b_2)}{b_1 <: b_2}$$
 subBase

4.
$$\frac{T_1' <: T_1 \qquad T_2 <: T_2'}{T_1 \to T_2 <: T_{21} \to T_2'}$$
 arrow

5.
$$\frac{T_1 <: T_2 \qquad T_2 <: T_3}{T_1 <: T_3}$$
 transitive

3.3 Examples

Type checking is complete for annotated simply typed lambda calculus with subtypes. This means the answer is **Yes** if the judgment can be derived, and **No** if the judgment can not be derived. The Unknown answer is never returned because annotated simple types are decidable.

Here are some examples tested by the **TypeChecker**. A red rule in the derivation tree mean its judgment is not derivable.

1. Valid variable rule

$$x: T \vdash x: T$$
 var

2. Invalid variable rule

$$\overline{\cdot \vdash x : T}$$
 invalid var

3. Lambda & application rules

$$\frac{x:(T1 \rightarrow T2), y:T1 \vdash x:(T1 \rightarrow T2)}{x:(T1 \rightarrow T2), y:T1 \vdash x:(T1 \rightarrow T2), y:T1 \vdash (x y)[T1]:T2} \text{app}$$

$$\frac{x:(T1 \rightarrow T2), y:T1 \vdash (x y)[T1]:T2}{x:(T1 \rightarrow T2) \vdash \lambda y.(x y)[T1]:(T1 \rightarrow T2)} \lambda$$

$$\frac{\cdot \vdash \lambda x.\lambda y.(x y)[T1]:((T1 \rightarrow T2) \rightarrow (T1 \rightarrow T2))}{\lambda}$$

- 4. Direct Subtyping
 - (a) Valid subtyping

$$\frac{x:bool \vdash x:bool}{x:bool \vdash x:bool} \text{ var} \quad \frac{\text{SubBase}(bool,int)}{bool <: int} \text{ subsumption}$$

$$x:bool \vdash x:int$$

(b) Invalid subtyping

5. Transitive subtyping

$$\frac{ \frac{\text{SubBase}(bool, int)}{bool <: int} \text{ subBase}}{x: bool \vdash x: bool} \frac{\frac{\text{SubBase}(bool, int)}{bool <: int} \text{ subBase}}{bool <: double}_{\text{subsumption}} \text{ transitive}$$

6. Transitive subtyping

| | ${\bf SubBase}(bool,int)$ | l D | ${\bf SubBase}(int, quotie$ | | | |
|--------------------------|---------------------------|----------------|-----------------------------|--------------------|---------------------------|-------------------------|
| | bool <: int | - subBase - | int <: quotient | subBase transitive | SubBase(quotient, double) | $\frac{ble)}{}$ subBase |
| ———var | | bool <: quotie | ent | transitive | quotient <: double | - transitive |
| $x:bool \vdash x:bool$ | | | transitive | | | |
| $x:bool \vdash x:double$ | | | | subsumption | | |

7. Arrow subtyping (subBase)

$$\frac{x:(int \to bool) \vdash x:(int \to bool)}{x:(int \to bool)} \text{ var } \frac{\frac{\text{SubBase}(bool,int)}{bool <: int} \text{ subBase}}{(int \to bool) <: (bool \to int)} \frac{\text{SubBase}(bool,int)}{bool <: int} \text{ arrow}}{(int \to bool) <: (bool \to int)} \text{ subsumption}$$

8. Arrow subtyping (reflexive, subBase)

$$\frac{x: (int \rightarrow bool) \vdash x: (int \rightarrow bool)}{x: (int \rightarrow bool)} \overset{\text{var}}{\underbrace{int <: int}} \overset{\text{feflexive}}{\underbrace{bool <: int}} \overset{\text{SubBase}(bool, int)}{\underbrace{bool <: int}} \overset{\text{subBase}}{\text{arrow}} \\ x: (int \rightarrow bool) \vdash x: (int \rightarrow int)$$
 subsumption

9. Arrow subtyping (reflexive, subBase)

$$\frac{SubBase(bool,int)}{bool <: int} subBase \frac{T <: T}{T} reflexive arrow \\ \frac{x: (int \to T), y: bool \vdash x: (int \to T)}{x: (int \to T), y: bool \vdash x: (bool \to T)} subsumption \frac{x: (int \to T), y: bool \vdash x: (bool \to T)}{x: (int \to T), y: bool \vdash x: (bool \to T)} var \\ \frac{x: (int \to T), y: bool \vdash x: (bool \to T)}{x: (int \to T) \vdash \lambda y. (x y)[bool] : (bool \to T)} \lambda \\ \frac{x: (int \to T) \vdash \lambda y. (x y)[bool] : (bool \to T)}{\vdash \lambda x. \lambda y. (x y)[bool] : ((int \to T) \to (bool \to T))} \lambda$$

10. Arrow subtyping (invalid)

$$\frac{\bot}{int <: bool} \text{ invalid} \quad \frac{\text{SubBase}(bool, int)}{bool <: int} \text{ subBase} \\ \frac{x : (bool \to bool) \vdash x : (bool \to bool)}{x : (bool \to bool) \vdash x : (int \to int)} \text{ subsumption}$$

11. Arrow subtyping (Invalid)

$$\frac{1}{x:(bool \rightarrow T), y:bool \vdash x:(bool \rightarrow T)} \text{ var } \frac{1}{\text{invalid}} \frac{1}{T < : T} \text{ reflexive arrow} \\ \frac{x:(bool \rightarrow T), y:bool \vdash x:(bool \rightarrow T)}{\text{total on of the problem}} \frac{1}{\text{total on of the pr$$

3.4 Delaying applying the subsumption rule

The sumbumption rule can be delayed until the term in the judgment is a variable. This simplifies the code since there is only one rule to be applied for application and λ abstraction terms. Here is the correctness proof using the subsumption rule:

$$\frac{\Gamma \vdash t : T \qquad T <: T'}{\Gamma \vdash t : T'} \text{ subsumption}$$

Proof. By induction hypothesis on the structure of the term t:

1. Case t = x: trivial since t is a variable.

$$\frac{\Gamma \vdash x : T \qquad T <: T'}{\Gamma \vdash x : T'} \text{ subsumption}$$

2. Case $t = t_1 t_2$: assume $T_2 <: T'_2$ and the following derivation:

$$\frac{\Gamma \vdash t_1 : T_1 \to T_2 \qquad \Gamma \vdash t_2 : T_1}{\Gamma \vdash (t_1 t_2)[T_1] : T_2} \text{ app} \qquad T_2 <: T_2'$$

$$\Gamma \vdash (t_1 t_2)[T_1] : T_2' \qquad \text{subsumption}$$

We can get a different derivation tree where the sumbsumption rule is delayed:

Using the induction hypothesis, the subsumption rule can be delayed until variable terms are generated in the derivations of $\Gamma \vdash t_1 : T_1 \to T_2'$ and $\Gamma \vdash t_2 : T_1$.

3. Case $t = \lambda x.t'$: assume $T_1 \to T_2 <: T_1' \to T_2'$ and the following derivation:

$$\frac{\Gamma, x: T_1 \vdash t': T_2}{\Gamma \vdash \lambda x. t': T_1 \to T_2} \lambda \qquad \frac{T_1' <: T_1 \qquad T_2 <: T_2'}{T_1 \to T_2 <: T_1' \to T_2'} \text{ arrow}$$

$$\Gamma \vdash \lambda x. t': T_1' \to T_2'$$
subsumption

Alternatively we can derive:

$$\frac{\Gamma, x: T_1' \vdash t': T_2 \qquad T_2 <: T_2'}{\frac{\Gamma, x: T_1' \vdash t': T_2'}{\Gamma \vdash \lambda x. t': T_1' \rightarrow T_2'}} \text{subsumption}$$

Assuming $T'_1 <: T_1$ is derivable, then by the subsumption rule:

$$\frac{\overline{\Gamma, x: T_1' \vdash x: T_1'} \quad \text{var} \quad T_1' <: T_1}{\Gamma, x: T_1' \vdash x: T_1} \text{ subsumption}$$

Therefore if $\Gamma, x: T_1 \vdash t': T_2$ is derivable, then $\Gamma, x: T_1' \vdash t': T_2$ is also derivable which concludes the proof.

Remark. If $\Gamma, x : T_1' \vdash t' : T_2$, it is not always true that $\Gamma, x : T_1 \vdash t' : T_2$ where $T_1' <: T_1$.

A counter example would be

$$x:bool \vdash x:int \Rightarrow x:double \vdash x:int, bool <: int <: double$$

However

$$x: int \vdash x: double \Rightarrow x: bool \vdash x: double, bool <: int <: double$$

Therefore in the following example, using the subsumption rule first would fail and slow the type checker because it needs to backtrack and check the λ rule. However using the λ rule first would prove the derivation and it is faster.

$$\frac{x: double \vdash x: int}{\Gamma \vdash \lambda x. x: double \rightarrow int} \lambda \frac{\frac{SubBase(bool, double)}{bool <: double} \text{ subBase}}{\frac{bool <: double}{double \rightarrow int} \text{ subsumption}} \frac{int <: int}{arrow} \text{ arrow}}{\text{arrow}} \frac{\Gamma \vdash \lambda x. x: bool \rightarrow int}{\Gamma, x: bool \vdash x: bool} \frac{SubBase(bool, int)}{bool <: int} \text{ subsumption}}{\frac{\Gamma, x: bool \vdash x: bool}{\Gamma \vdash \lambda x. x: bool \rightarrow int}} \lambda$$

4 System F

Annotated system F terms, and types recognized by the parser have the form:

$$t ::= x \mid (t_1 t_2)[T] \mid \lambda x.t \mid t [[T]]$$
$$T ::= X \mid T_1 \to T_2 \mid \forall X.T$$

The annotation [T] in the application term $(t_1t_2)[T]$ is used to remove the non-determinism in the application rule. The annotation [[T]] in the term t[[T]] is used to remove the non-determinism in the elimination rule. Unlike the application annotation, the elimination annotation is not required in the parser which makes the **TypeChecker** incomplete. Whenever the **TypeChecker** needs the elimination annotation and it is not provided, it returns the current derivation tree with answer **Unknown**.

4.1 Rules

1. If X is not free in Γ

$$\frac{\Gamma \vdash t : T}{\Gamma \vdash t : \forall X.T} \text{ introduction}$$

2. If X is free in Γ , choose X_i such that X_i is not free in Γ

$$\frac{\Gamma \vdash t : [X_i/X]T}{\Gamma \vdash t : \forall X_i.[X_i/X]T} \frac{\text{introduction}}{\Gamma \vdash t : \forall X.T}$$
renaming

3. Elimination rule with annotation

$$\frac{\Gamma \vdash t : \forall X.[X/T']T}{\Gamma \vdash t \; [[T']] : T} \; \text{elimination}$$

4.2 Examples

Here are some examples tested by the **TypeChecker**. A red rule in the derivation tree mean its judgment is not derivable. A blue rule means the **TypeChecker** needs more annotation to continue the type checking.

1. Same type variable

$$\overline{x: \forall X.X \vdash x: \forall X.X}$$
 var

2. Different type variables

$$x: \forall X.X \vdash x: \forall Y.Y$$
 var

3. Arrows

$$x: \forall X.(X \to X) \vdash x: \forall Y.(Y \to Y)$$
 var

4. Y is free in the typing context and the term type

$$x: \forall X.(X \to Y) \vdash x: \forall Z.(Z \to Y)$$
 var

5. Y is free in the typing context

$$x: \forall X.(X \to Y) \vdash x: \forall Y.(Y \to Y)$$
 invalid var

6. Y is free in the typing context, and Z is free in the term type

$$x: \forall X.(X \to Y) \vdash x: \forall Y.(Y \to Z)$$
 invalid var

7. Elimination annotation

$$\frac{ \overline{x: \forall X.X \vdash x: \forall X1.X1}}{x: \forall X.X \vdash x[Y]: Y}^{\text{var}}$$
 elimination

8. Elimination annotation with arrow

$$\frac{x : \forall X.X \vdash x : \forall X1.X1^{\text{var}}}{x : \forall X.X \vdash x[(Y \to Y)] : (Y \to Y)} \text{ elimination}$$

9. Nested elimination annotation

$$\frac{ x : \forall X.X \vdash x : \forall X1.X1}{x : \forall X.X \vdash x : \forall X1.X1} \text{ var} \\ \frac{ x : \forall X.X \vdash x[[(\forall X.X \to \forall X.X)]][(\forall X.X \to \forall X.X)] : (\forall X.X \to \forall X.X)}{x : \forall X.X \vdash (x[[(\forall X.X \to \forall X.X)]] \ x)[\forall X.X] : \forall X.X} \text{ app}$$

10. Application or introduction non-determinism: The **TypeChecker** attempts applications first. When it fails, it backtracks and attempts the introduction rule as shown in the next example.

7

$$\frac{\frac{\bot}{T_1 <: T_1} \text{ reflexive } \frac{\bot}{T_2 <: \forall X.T_2} \text{ invalid}}{(T_1 \to T_2), y : T_1 \vdash x : (T_1 \to T_2)} \text{ var } \frac{(T_1 \to T_2) <: (T_1 \to \forall X.T_2)}{(T_1 \to T_2) <: (T_1 \to \forall X.T_2)} \text{ subsumption } \frac{x : (T_1 \to T_2), y : T_1 \vdash x : (T_1 \to \forall X.T_2)}{x : (T_1 \to T_2), y : T_1 \vdash (x, y)[T_1] : \forall X.T_2} \text{ app}$$

11. Introduction branch

$$\frac{x: (T_1 \to T_2), y: T_1 \vdash x: (T_1 \to T_2)}{x: (T_1 \to T_2), y: T_1 \vdash (x y)[T_1]: T_2} \stackrel{\text{var}}{\underset{\text{app}}{\underbrace{x: (T_1 \to T_2), y: T_1 \vdash (x y)[T_1]: T_2}}} \xrightarrow{\text{introduction}}$$

12. Application or elimination non-determinism

$$\frac{x: (T_1 \rightarrow \forall X.T_2), y: T_1 \vdash x: (T_1 \rightarrow \forall X.T_2)}{x: (T_1 \rightarrow \forall X.T_2), y: T_1 \vdash x: (T_1 \rightarrow \forall X.T_2)} \text{ var } \frac{\frac{\bot}{T_1 <: T_1} \text{ reflexive } \frac{\bot}{\forall X.T_2 <: T_2} \text{ invalid arrow}}{(T_1 \rightarrow \forall X.T_2) <: (T_1 \rightarrow T_2)} \text{ subsumption } \frac{x: (T_1 \rightarrow \forall X.T_2), y: T_1 \vdash x: (T_1 \rightarrow T_2)}{x: (T_1 \rightarrow \forall X.T_2), y: T_1 \vdash (x y)[T_1][[Y]]: T_2} \text{ var approximation}$$

13. Elimination branch

$$\frac{x: (T_1 \to \forall X.T_2), y: T_1 \vdash x: (T_1 \to \forall X_1.T_2)}{x: (T_1 \to \forall X.T_2), y: T_1 \vdash x: T_1} \operatorname{app}_{\text{app}} \frac{x: (T_1 \to \forall X.T_2), y: T_1 \vdash (x y)[T_1]: \forall X_1.T_2}{x: (T_1 \to \forall X.T_2), y: T_1 \vdash (x y)[T_1][[Y]]: T_2} \operatorname{elimination}_{\text{elimination}}$$

14. Zero

$$\frac{\overline{z:X,s:(X\to X)\vdash z:X}^{\text{var}}}{s:(X\to X)\vdash(\lambda z.z):(X\to X)}\lambda$$

$$\frac{\cdot\vdash(\lambda s.(\lambda z.z)):((X\to X)\to(X\to X))}{\cdot\vdash(\lambda s.(\lambda z.z)):\forall X.((X\to X)\to(X\to X))}^{\text{introduction}}$$

15. Zero with free variable X

$$\frac{\frac{y:X,z:X_2,s:(X_2\to X_2)\vdash z:X_2}{y:X,s:(X_2\to X_2)\vdash (\lambda z.z):(X_2\to X_2)}\lambda}{y:X\vdash (\lambda s.(\lambda z.z)):((X_2\to X_2)\to (X_2\to X_2))}\lambda}{y:X\vdash (\lambda s.(\lambda z.z)):\forall X_2.((X_2\to X_2)\to (X_2\to X_2))}_{\text{renaming}} \xrightarrow{\text{renaming}}$$

Successor missing annotation

| 4 02 | ימו | 444 | | | | | |
|--|--|---|--|--|--|--|---|
| | $n: \forall X. ((X \rightarrow X) \rightarrow (X \rightarrow X)), z: X, s: (X \rightarrow X) \vdash z: X$ | d d | 445 | | | | |
| $ \text{var} = \frac{\text{var}((X \rightarrow X) + (X \rightarrow X)), z : (X \rightarrow X) + n : ((X \rightarrow X) + (X \rightarrow X))}{\text{orn}} $ | $n: \forall X. ((X \rightarrow X) + (X \rightarrow X)); z: X, z: (X \rightarrow X) + (n z) [(X \rightarrow X)]; (X \rightarrow X)$ | $n: \forall X. ((X \rightarrow X) \rightarrow (X \rightarrow X)), z: X, z: (X \rightarrow X) \vdash ((n \rightarrow)[(X \rightarrow X)] z)[X]: X$ | $n: \forall X.((X \rightarrow X) \rightarrow (X \rightarrow X)), z: \chi, s: (X \rightarrow X) \vdash (s: (n:s)[(X \rightarrow X)] \Rightarrow [s: (X \rightarrow X)$ | $n: \forall X. ((X \rightarrow X) \rightarrow (X \rightarrow X)), s: (X \rightarrow X) \vdash (\lambda a. (s \cdot (s \cdot a)((X \rightarrow X)) \mid S)[X])[X]) : (X \rightarrow X)$ | $n: \forall X. ((X \rightarrow X) \rightarrow (X \rightarrow X)) + (\lambda x. (\lambda x. (x \circ \beta)[(X \rightarrow X)] \pm)[X])[X])) : ((X \rightarrow X) \rightarrow (X \rightarrow X))$ | $n: \forall X.((X \to X) \to (X \to X)) \vdash (\lambda x.(\lambda x.(x ((x \to X)) \to (X \to X)) \to (X \to X)) \vdash (X \to X)) \vdash (X \to X)) \vdash (X \to X) \vdash (X \to X)) \vdash (X \to X) \vdash (X \to $ | $ + \left(\Lambda n_*(\lambda *_*(x *_s((x *_s)[(X \to X)] *_s[X])(X)) \right) \cdot (\forall X \cdot ((x \to X) \to (X \to X)) \to \forall X \cdot ((X \to X) \to (X \to X))) $ |
| | N CAR | $n: \forall X.((X \rightarrow X) \rightarrow (X \rightarrow X)), z: X, s: (X \rightarrow X) \vdash s: (X \rightarrow X)$ | | | | | |

Successor well annotated

| $\frac{\text{VBF}}{\text{ATR}} \times X \times$ | ddn ——————————————————————————————————— | Ada I | | | | |
|---|---|---|--|---|--|--|
| $ \underset{n+\forall X((X+X)\to(X+X); n+X, n+(X+X) \to (X+X))}{n+\forall X((X+X)\to(X+X) \to (X+X); n+(X+X)} = \underset{n+\forall X((X+X)\to(X+X); n+(X+X)}{\text{elimination}} $ | $n: \forall X((X \rightarrow X) \rightarrow (X \rightarrow X)), z: X, z: (X \rightarrow X)) + ((\alpha[[X]] \circ [[X \rightarrow X)] : X)$ | $n: \forall X.((X \rightarrow X) \rightarrow (X \rightarrow X)), z: X, z: (X \rightarrow X) \vdash (s.(n[X]] \Rightarrow l[(X \rightarrow X)] \Rightarrow l[(X)](X) \mid X$ | $n: \forall X. ((X \rightarrow X)), s: (X \rightarrow X) \vdash (\Delta_{x}(s ([X][X]] s)[(X \rightarrow X)] : s)[X][X]): (X \rightarrow X)$ | $n: \forall X : (X \rightarrow X) \rightarrow (X \rightarrow X)) \vdash (\lambda_{x} : (\lambda_{x} : (a \cdot (n [X] : s) (x \rightarrow X)] : a) X) X) X) \mid (X \rightarrow X) \rightarrow (X \rightarrow X)$ | $n: \forall x . (X \rightarrow X) + (X \rightarrow X)) + (\lambda x. (\lambda x. (x \cdot ($ | $+ (\lambda_m (\lambda_n (\lambda_n (\{x \in (x \in (x \in [X]], x)[x](X) \mid x)(X \setminus (X \neq X) \to (X + X)) + v_X ((X + X) \to (X \to X))) \xrightarrow{\lambda} \lambda_m (X \to X) = \lambda_m (X \to X) + v_X $ |
| | $n: \forall X. ((X \rightarrow X) \rightarrow (X \rightarrow X)), z: X, s: (X \rightarrow X) \vdash s: (X \rightarrow X)$ | | | | | |